



UNIVERSIDADE D
COIMBRA

Inês Nobre Cortesão Júdice Pargana

**SIMULAÇÃO, TESTES DE VALIDADE DE SINTOMAS E *DESIGN*
DE GRUPOS CONHECIDOS: UMA REVISÃO DE LITERATURA
EM CONTEXTO FORENSE**

Dissertação no âmbito do Mestrado em Psicologia Clínica e da Saúde, subespecialização em
Psicologia Forense, orientada pelo Professor Doutor Mário R. Simões.
e apresentada à Faculdade de Psicologia e Ciências da Educação da Universidade de Coimbra.

Outubro de 2020

**Simulação, Testes de Validade de Sintomas e *Design* de Grupos
Conhecidos: Uma revisão de literatura em Contexto Forense**

Resumo: O problema da validade de sintomas é transversal aos vários contextos de avaliação (neuro)psicológica. Por isso, os protocolos de avaliação devem incluir o exame de vários estilos de resposta, no sentido de examinar a credibilidade dos resultados. Neste âmbito, a avaliação da simulação ou exagero de sintomas constitui uma tarefa fundamental que deve ser realizada através de instrumentos específicos, denominados Testes de Validade de Sintomas (TVS).

Comparativamente a outros contextos de avaliação, a prevalência da simulação de sintomas é mais elevada em contextos forenses sendo, por isso, necessária a validação de medidas específicas que permitam responder a tais necessidades. O meio prisional constitui um ambiente adverso no qual existe maior probabilidade de simulação de sintomas na resposta aos instrumentos e técnicas de avaliação. Os ganhos secundários, decorrentes de tais respostas enviesadas, podem assegurar melhores condições ou mesmo redução do cumprimento da pena. Por isso, são indispensáveis instrumentos específicos empiricamente validados que garantam a deteção da simulação e de outros estilos de resposta.

Neste plano, o *design* de grupos conhecidos constitui uma metodologia com reconhecidas potencialidades, sobretudo ao nível da validade externa. Através da análise deste *design*, o presente trabalho procura reunir informação proveniente da investigação empírica de natureza psicométrica sobre a simulação de sintomas em vários contextos forenses, no âmbito do desenvolvimento e validação de Testes de Validade de Sintomas.

Palavras chave: avaliação psicológica forense; simulação de sintomas; contexto forense; *design* de grupos conhecidos.

Malingering, Validity Tests Symptoms and Known-Groups Design: A review of the literature in a Forensic Context

Abstract: The problem of symptom validity is transversal to the various contexts of (neuro) psychological assessment. For this reason, the assessment protocols must include the examination of several response styles, in order to examine the credibility of the results. In this context, the assessment of the malingering or exaggeration of symptoms is a fundamental task that must be performed through specific instruments, called Symptom Validity Tests.

Compared to other assessment contexts, the prevalence of malingering symptoms is higher in forensic contexts and, therefore, it is necessary to validate specific measures in order to respond to such evaluation needs. The prison context is an adverse environment, in which there is a greater likelihood of symptom malingering in response to assessment instruments and techniques. Secondary gains, resulting from such skewed responses, can ensure better conditions or even a reduction in the sentence. For this reason, empirically validated instruments are essential to guarantee the detection of malingering and other response styles.

In this plan, the known groups design is a methodology with recognized potential, especially when it comes to external validity. Through the analysis of this specific design, the present work seeks to gather information from empirical investigation of a psychometric nature about malingering or exaggeration of symptoms in various forensic contexts, within the scope of the development and validation of Symptom Validity Tests.

Key Words: forensic psychological assessment; malingering; forensic context; known-groups design.

Agradecimentos

Depois daquele que foi, certamente, o ano mais difícil e desafiante da minha vida, não posso deixar de agradecer:

Ao Professor Mário Simões, pela orientação e pelos conhecimentos partilhados ao longo de um percurso que teve os seus obstáculos e que tantas adaptações exigiu.

À Dra. Isabel Teixeira Dias, por toda a disponibilidade, partilha e apoio, que tornaram o meu estágio curricular na experiência mais gratificante do meu percurso académico.

À minha mãe, sem a qual este trabalho dificilmente teria sido concluído – pela força, pelo colo e, sobretudo, pelo amor incondicional.

Ao meu pai, por termos conseguido encontrar de novo o nosso caminho.

À minha avó, que é, indubitavelmente, a pessoa mais bonita que encontrarei na vida; e ao meu avô, por ser o meu maior exemplo de coragem.

Às minhas meninas – Joana, Eva, Francisca e Ana, por cuidarem sempre de mim e me mostrarem todos os dias quão especial é a família que escolhi.

Um sincero Obrigada.

Índice

Introdução	1
I – Avaliação da Validade de Sintomas – Estilos de Resposta.....	2
1.1 – Simulação e Exagero de Sintomas	4
II – Avaliação da Simulação e outros estilos de resposta	7
2.1 Testes de Validade de Desempenho (TVD)	14
2.2 Testes de Validade de Sintomas (TVS).....	15
III – <i>Designs</i> de validação: o <i>design</i> de grupos conhecidos.....	19
IV – Avaliação da Simulação em Contexto Forense	21
V – Questões éticas na avaliação da Simulação de Sintomas	26
VI - <i>Self-Report Symptom Inventory</i> (SRSI) em contexto prisional: Proposta de estudo com <i>design</i> de grupos conhecidos	29
VII – Conclusão.....	31
Bibliografia	33

Introdução

As avaliações (neuro)psicológicas baseiam-se no pressuposto de que os resultados do sujeito deverão refletir o seu funcionamento global e específico, uma vez que se considera que as respostas aos instrumentos terão sido dadas com base no esforço e motivação máximos. Se essa intenção for, por alguma razão, comprometida e o paciente alterar as suas respostas, o perfil psicológico obtido será, simplesmente, inválido (Rogers & Cruise, 1998; Merten & Merckelbach, 2013).

A necessidade da inclusão da avaliação da validade de sintomas nos protocolos tem sido reconhecida, tanto em contextos clínicos, como forenses (Martin et al., 2015). Contudo, importa realçar que as prevalências diferem quanto ao contexto, sendo expectável que a percentagem de relatos deturpados seja manifestamente superior em contextos forenses (Bush, Ruff, Troster, Barth, Kofflet, Pliskin, Reynolds, & Silver, 2005; Cornell & Hawk, 1989; Iverson, 2003; Simões, 2012). De facto, a simulação de sintomas constitui a maior ameaça à validade das avaliações psicológicas em contextos forenses (Giger & Merten, 2019; Iverson, 2003; Pollock, Quigley, Worley, & Bashford, 1997; Rogers, 2018).

A avaliação da validade de sintomas deverá basear-se em várias medidas e estratégias no sentido de aumentar a probabilidade da sua correta identificação (Larrabee, 2008). Nos últimos anos, têm surgido vários instrumentos específicos para a avaliação deste fenómeno que, integrados nos protocolos de avaliação, contribuem para garantir a credibilidade dos resultados. Algumas destas medidas encontram-se já validadas para os diferentes contextos avaliativos, havendo necessidade da validação de mais instrumentos para o contexto forense e prisional, especificamente. Apesar de não ser tão explorado, comparativamente a outros métodos, o *design* de grupos conhecidos assume potencialidades que sustentam a sua utilização na validação destes instrumentos (Rogers, 2018).

O presente trabalho pretende explorar as potencialidades do *design* de grupos conhecidos na avaliação da simulação de sintomas, particularmente em contexto forense, onde os ganhos secundários se mostram mais prevalentes. Adicionalmente, é proposto um estudo empírico que utiliza o *design* de grupos conhecidos no contexto da necessidade de validação de uma nova medida de autorrelato: o *Self-Report Symptom Inventory* (SRSI), com estudos recentes em Portugal. Por fim, e considerando que a investigação recente tem sido bastante crítica no que toca a este campo avaliativo, serão exploradas algumas questões éticas que deverão ser consideradas aquando da avaliação da credibilidade dos relatos.

I – Avaliação da Validade de Sintomas – Estilos de Resposta

A omissão ou alteração de informação está implícita na maioria das relações interpessoais, independentemente do contexto (Rogers, 2018a). As relações terapêuticas não são, de todo, exceção. Apesar do indivíduo poder estar comprometido e motivado em relação ao processo avaliativo/terapêutico, tal não exclui a possibilidade de este omitir ou deturpar informação relativa a si próprio. Tal pode acontecer tanto no sentido de procurar passar uma imagem mais favorável de si, amplificando características positivas e/ou minimizando ou negando sintomatologia por receio de julgamento social, ou pode suceder-se que queira chamar a atenção do terapeuta para um dado aspeto, realçando dificuldades através de um exagero na comunicação das suas fragilidades e sintomas (Rogers, 2018a). Neste sentido, não poderá ser modestamente assumido que o sujeito deposite no protocolo de avaliação todo o esforço e empenho que seria expectável. Há que considerar que a validade dos resultados poderá ser posta em causa devido a alguns fatores que afetam o compromisso do sujeito para com o processo avaliativo, tais como a voluntariedade da participação, fatores situacionais (ansiedade, fadiga,...), presença de algum tipo de perturbação mental que afete o desempenho, ou a evidência de algum estilo de resposta específico, como o esforço reduzido, o fingimento, a dissimulação, ou a simulação de sintomas ou défices (Simões, 2006; 2012).

Para que os resultados e as conclusões inferidas de um protocolo de avaliação sejam válidas é imperativo que exista empenho, motivação, esforço e honestidade na resposta aos instrumentos. Caso tal não aconteça, a interpretação dos perfis de resultados é posta em causa, uma vez que não traduzirá o real funcionamento do sujeito (Dandachi-FitzGerald, Duits, Leentjens, Verhey & Ponds, 2020; Larrabee, 2007; Rogers, 2018a; Tombaugh, 1996). De facto, nas últimas décadas, a investigação no âmbito da validade de sintomas tem aumentado exponencialmente (Bigler, 2012; Carr, 2015; Rogers, 2018a) e têm sido identificados diversos estilos de resposta que os indivíduos podem adotar perante o protocolo de avaliação. Nos instrumentos de autorrelato, é possível encontrar uma variedade de estilos de respostas que podem enviesar os resultados, uma vez que inviabilizam o perfil do sujeito. Na década de 80, Paulhus (1984; citado em McGrath, Mitchell, Kim & Hough, 2010) propõe que os enviesamentos de resposta ocorrem tendencialmente com base em dois tipos de motivações: a primeira, no sentido de enganar o examinador, ocorre maioritariamente em situações de recrutamento (em que o indivíduo procura transmitir uma imagem favorável de si); e a segunda em contextos de pedidos de baixa médica, incapacidades, atenuações de culpa (baseadas na tentativa de produzir ou enaltecer fragilidades ou incapacidades). Um enviesamento de resposta representa a tendência consistente para responder de forma inadequada perante um dado instrumento, baseando-se a resposta em fatores que não traduzem as verdadeiras crenças, pensamentos e comportamentos do indivíduo (McGrath et al., 2010). Os estilos de resposta mais facilmente compreendidos, considerando a investigação científica que sobre eles tem recaído, são o esforço reduzido ou insuficiente, a defensividade, a deseabilidade social, ou

a simulação de sintomas (Bush et al., 2005; McGrath et al., 2010; Simões, 2012; Rogers, 2018a).

O *esforço reduzido* (também denominado *esforço insuficiente*) representa um estilo de resposta em que o o sujeito não está ativamente envolvido no processo de avaliação, respondendo de forma pouco esforçada, inconsistente ou intencionalmente aleatória, refletindo um esforço deliberadamente fraco na resposta aos instrumentos de avaliação (Heilburn, 1992). A *desejabilidade social* representa a tendência para transparecer uma imagem favorável de si com base em comportamentos/attitudes que o sujeito considera cultural e moralmente aceitáveis, rejeitando a presença de atitudes/comportamentos com valores socialmente indesejáveis (Rogers, 2018a). Por sua vez, a *defensividade* representa a denegação ou minimização consciente e deliberada de dificuldades ou incapacidades (Heilburn, 1992). No extremo oposto a este conceito, está a *simulação* ou *exagero de sintomas*, que se define como a produção intencional de sintomas estimulada por motivações externas (APA, 2014; Rogers, 2018a). Esta questão será mais extensivamente analisada na rúbrica seguinte.

No entanto, existem alguns estilos de resposta que não são, ainda, bem compreendidos pela comunidade científica (McGrath et al., 2010; Rogers, 2018a), apesar de a sua existência ser explorada e estar confirmada. É o caso da *tendência para respostas inconsistentes ou irrelevantes*, que se manifesta quando o indivíduo não está suficientemente comprometido com o processo de avaliação e, por isso, seleciona respostas aleatoriamente (Rogers, 2018a). Outro exemplo, a *aquiescência* representa a tendência para responder afirmativamente, não devendo este conceito ser confundido com a desejabilidade social que implica outro propósito na seleção das respostas. Contrariamente a este conceito, o *negativismo* assume-se como a tendência para selecionar respostas formuladas no sentido negativo. Quando empregue em escalas de psicopatologia, por exemplo, os sujeitos procuram selecionar itens que não transpareçam sintomatologia psicológica. É ainda possível verificar a existência estilos de resposta que se relacionam com a estrutura da escala: tendência para a seleção de respostas situadas no extremo da escala (respostas extremas) ou para a seleção de itens mais no meio da escala (respostas neutras). Alguns indivíduos podem adotar vários estilos de resposta ao longo de um mesmo processo de avaliação, assumindo um *estilo de resposta híbrido* (Rogers, 2018a; Lynch, 2004).

As avaliações (neuro)psicológicas deverão incluir medidas que visem a avaliação dos múltiplos estilos de resposta, de forma a assegurar a validade do protocolo (Iverson, 2000; Lilienfeld, Thames & Watts, 2013; Lynch, 2004). Quando não é possível a inclusão de vários instrumentos neste sentido, é da responsabilidade do examinador a justificação das implicações que poderão refletir-se na interpretação do protocolo (Heilbronner et al., 2009).

1.1 – Simulação e Exagero de Sintomas

A conceitualização e os fatores que conduzem ao diagnóstico da simulação têm vindo a ser desenvolvidos e aperfeiçoados, ao longo das últimas décadas. De acordo com o manual DSM-5 (American Psychiatric Association [APA], 2014), a simulação é definida como a produção intencional de sintomas físicos ou psicológicos de forma falsa ou exagerada, motivada por incentivos externos (APA, 2014). O manual não conceitualiza a simulação enquanto patologia formal, incluindo-a na secção das condições que deverão ser foco de atenção clínica, uma vez que poderão afetar o diagnóstico, curso, prognóstico, ou tratamento de uma dada doença. De uma forma mais específica, o DSM-5 (APA, 2014) inclui, na definição de simulação, três componentes fundamentais da condição: a produção/exagero de sintomas, a motivação e a intenção consciente do sujeito para simular, e a presença de incentivos externos. O comportamento de simulação resulta de uma análise de custo-benefício, de acordo com a qual o sujeito escolhe a produção ou exagero dos sintomas como meio de atingir os seus objetivos pessoais. Para que o diagnóstico de simulação seja válido, deverá haver fabricação ou evidente exagero de múltiplos sintomas, atendendo a que a presença de apenas pequenos exageros não deverá conduzir ao seu diagnóstico. Para além disto, importa considerar que a presença de incentivos externos não exclui a possibilidade de coocorrência de motivações internas para simular (Simões, 2005; Rogers, 2018). Assim, torna-se fundamental que as pontuações dos instrumentos que procuram avaliar esta condição tenham qualidades psicométricas, como a sensibilidade e a especificidade. A simulação pode manifestar-se em vários padrões, nomeadamente sob a forma de falsificação ou comunicação de défices não existentes (“simulação pura”), que é um fenómeno mais raro; através do exagero de sintomatologia real experienciada pelo sujeito (“simulação parcial”), que é mais comum encontrar em contexto avaliativo; através da comunicação da manutenção ou agravamento de sintomas previamente vivenciados; ou através da incorreta atribuição de sintomas, falsamente imputados a uma causa não diretamente relacionada (Heubrock & Petermann, 1998; Rogers, 2018).

Várias teorias têm procurado explicar a simulação de sintomas. A teoria que reúne maior aceitação é conhecida por modelo de adaptação, proposto por Rogers (1990, 2018). Este modelo apresenta uma explicação simples para a simulação de psicopatologia, tendo sido a partir destas ideias que a simulação passou a ser vista como um estilo de resposta e não como um diagnóstico em si (Rogers, 1990; Pollock, et al., 1997). O modelo adaptativo propõe que a simulação de sintomas se baseia numa análise custo-benefício perante a escolha entre simular e outras alternativas, constituindo este comportamento ou estilo de resposta uma reação normativa perante uma situação difícil (Rogers, 1990). De uma forma geral, o modelo postula que para que haja produção/exagero intencional de sintomas, três condições deverão estar presentes: o indivíduo percebe o contexto de avaliação como adverso; percebe que poderá ter alguma coisa a perder ao transparecer a verdadeira imagem de si, e alguma coisa a ganhar se adulterar essa mesma imagem; e não consegue encontrar outro meio para atingir o objetivo

desejado. Se o modelo estiver correto, quanto maiores os ganhos secundários, maior a probabilidade de o sujeito enveredar por um perfil de respostas simuladas (McDermott, Dualan & Scott, 2013).

Existem dois outros modelos que procuram explicar o fenómeno da simulação, não reunindo, contudo, tanto consenso quanto o modelo da adaptação: o modelo patogénico e o modelo criminológico. O modelo patogénico está relacionado com o pensamento psicodinâmico, postulando que a motivação para simular advém de uma perturbação presente no indivíduo, i.e., o sujeito simulador não consegue controlar a tendência para produzir ou exagerar sintomas (Rogers, 1990, 2018). Apesar de não ser o modelo que melhor explica a simulação nas mais diversas valências (Rogers et al, 1998), algumas manifestações deste fenómeno poderão ser suportadas por esta visão (Vitacco & Rogers, 2005).

Por sua vez, o modelo criminológico encontra-se diretamente correlacionado com a visão apresentada pelo DSM. De acordo com este modelo, a simulação é um perfil deliberadamente escolhido por sujeitos com perturbação ou comportamentos antissociais (Rogers, 1990, 2018). No entanto, esta relação não deverá ser diretamente assumida uma vez que, em contextos forenses, a grande maioria dos indivíduos apresenta historial de comportamentos antissociais, não se verificando, contudo, que a totalidade dessa mesma maioria apresente um perfil de simulação. Por essa razão, de acordo com a comunidade científica, o modelo criminológico é tido como pouco específico e válido (Vitacco & Rogers, 2005). Mesmo em contexto prisional, o modelo de adaptação é tido como preferencial na interpretação da simulação, uma vez que a resposta simulada poderá dever-se a uma tentativa de o recluso atingir um dado objetivo (Vitacco & Rogers, 2005).

Independentemente do contexto, o diagnóstico da simulação deve atender a um conjunto de princípios que, não sendo óbvios, podem conduzir a algumas dificuldades na diferenciação perante outras perturbações/condições. Contrariamente à simulação, a Perturbação Factícia constitui-se como uma perturbação formalmente reconhecida pelo DSM-5 (APA, 2014). Na Perturbação Factícia o indivíduo assume um perfil de relato exagerado de sintomas com o propósito de assumir o papel de “doente”, sem intenção ou consciência dos ganhos secundários que daí poderão advir, isto é, o relato de sintomas prende-se com motivações inconscientes, determinadas pelo mero ganho primário de satisfação de uma necessidade psíquica interna (APA, 2014; Conroy & Kwarter, 2006; Gil et al., 2017; van Impelen, Merckelbach, Jelicic, Niesten & Campo, 2017). Importa sublinhar que, nesta perturbação, o comportamento de dissimulação está presente mesmo que não se verifiquem incentivos externos, apesar de poder ocorrer na presença deles. Tratando-se de uma perturbação formal, estes indivíduos necessitam de tratamento e acompanhamento psiquiátrico, experienciando elevados níveis de mal-estar e défices funcionais, sofrimento este não necessariamente presente num perfil de simulação. Na prática clínica a avaliação da motivação por detrás de um relato falseado ou exagerado de sintomas é tendencialmente difícil (Slick, Sherman & Iverson, 1999), pelo que a linha diferenciadora entre estas condições é muito ténue, estando sobretudo relacionada com a

motivação subjacente aos comportamentos de dissimulação (Seron, 2014; Yates et al., 2018). Independentemente da perturbação ou estilo de resposta em análise, o diagnóstico diferencial da simulação de sintomas exige que a avaliação do profissional seja baseada em critérios objetivos, incorporando indicadores suficientemente validados que corroborem a fiabilidade do diagnóstico (Heilbronner et al., 2009).

A emergente investigação associada à simulação de sintomas tem conduzido ao surgimento de algumas críticas à definição inicial apresentada pelo DSM-III, que se traduziram na necessidade de uma caracterização mais fundamentada deste fenómeno, tanto em termos teóricos, como empíricos (Simões, 2012). De acordo com a definição mais recente do DSM-5 (APA, 2014), a motivação é o construto central que deverá ser considerado aquando da avaliação deste problema. Contudo, a definição de simulação é muito mais clara na teoria do que na prática, uma vez que, em contexto real, é marcadamente mais difícil reconhecer e compreender as motivações que levam um dado a indivíduo a enveredar por este estilo resposta (Rogers, 2018). Assim, muitos investigadores debatem-se perante a concetualização da simulação numa perspetiva dicotómica (respostas genuínas ou respostas simuladas), ou enquanto fenómeno enquadrado num espectro contínuo (Carr, 2015). Neste sentido, importa salientar que a simulação não é um fenómeno extremo de “tudo-ou-nada”. De facto, são raras as situações de “simulação pura” (Rogers, 2018b), pelo que é pertinente a sua interpretação enquanto fenómeno constituído por vários níveis e, por isso, deve ser analisado num espectro contínuo e dimensional. Para além disto, importa considerar que na grande maioria das situações, os sujeitos não simulam sintomatologia em todas as áreas de avaliação. Não é expectável que a suspeita de simulação num determinado âmbito justifique eventuais valores em todas as áreas de funcionamento do indivíduo (Heilbronner et al., 2009). Com efeito, em termos clínicos, poderá ser mais gravoso para o paciente que não lhe seja feito o diagnóstico da patologia que realmente tem, do que diagnosticar alguém que não tem qualquer condição patológica, uma vez que no primeiro cenário estariam a ser negados cuidados de saúde a um doente (Barber-Rioja et al., 2016; Faust, 1995). É fundamental perceber que a simulação e a psicopatologia não são mutuamente exclusivas, podendo existir concomitantemente no indivíduo (Berry & Nelson, 2010), pelo que o examinador deverá considerar várias explicações para a presença de queixas excessivas (Heilbronner et al., 2009; Rogers, 2018). Para além destes aspetos, o DSM-5 propõe que os profissionais estejam mais atentos à presença de simulação caso se evidenciem duas ou mais das seguintes características: o facto de se tratar de um contexto forense, devido à maior probabilidade de ganhos secundários, a existência de notórias discrepâncias entre os dados processuais clínicos e o autorrelato individual, evidência de falta de cooperação e motivação por parte do examinado, e a presença de Perturbação Antissocial da Personalidade (APA, 2014). Este último aspeto tem vindo a ser alvo de bastantes críticas, dado que alguns investigadores afirmam que o fenómeno da simulação não se circunscreve a indivíduos perturbados, em circunstâncias adversas ou com comportamentos desviantes (Rogers, 1990).

No sentido de neutralizar as fragilidades encontradas na definição proposta pelo DSM-5, este problema tem vindo a ser explorado nos últimos anos. Contudo, considerando que o manual em causa baseia a prática clínica da maioria dos profissionais em todo o mundo, é possível que estes restrinjam os conhecimentos acerca da simulação unicamente a estas indicações (Rogers, 1990; Berry & Nelson, 2010; Rogers, 2018), o que poderá conduzir a classificações erróneas e ao aumento do número de falsos positivos (McDermott & Sokolov, 2009).

II – Avaliação da Simulação e outros estilos de resposta

Um dos objetivos principais da avaliação (neuro)psicológica deverá ser o evitamento do erro tipo I, de forma a que pacientes com efetiva psicopatologia não tenham negado o acesso a cuidados de saúde mental (Vitacco, Rogers, Gabel & Munizza, 2007). No entanto, se a avaliação da validade dos relatos se basear exclusivamente na intuição clínica do profissional, a investigação comprova que é relativamente fácil exagerar sintomatologia de forma a que tal seja indetetável. Assim, o desenvolvimento de instrumentos que visam a avaliação da validade de respostas e sintomas tem vindo a mostrar-se indispensável (Vitacco et al., 2007; Merten & Merckelbach, 2013).

A informação no âmbito da validade de respostas poderá ser retirada de várias fontes, ao longo de todo o processo avaliativo. Apesar de os examinadores não deverem limitar esta avaliação ao uso de instrumentos clássicos, algumas medidas neste âmbito poderão ser um bom ponto de partida no processo inicial de triagem, no sentido de determinar se o indivíduo beneficiará de uma avaliação mais focalizada e compreensiva deste campo (Edens, Poythress & Watkins-Clay, 2007). Alguns instrumentos de avaliação clássicos, com estudos portugueses, como a Wechsler Adult Intelligence Scale (WAIS-III) ou o Trail Making Test (TMT) são frequentemente considerados na avaliação da validade de desempenhos, uma vez que poderão fornecer importantes afirmações através das análises de perfil de respostas/desempenhos e do padrão de resultados (Mittenberg, Patton, Canyock & Condit, 2002; Wygant, Walls, Brothers & Berry, 2018). Também algumas medidas tradicionais de avaliação da personalidade, em fase de adaptação e validação para a população portuguesa, como o Minnesota Multiphasic Personality Inventory (MMPI-2) ou o Personality Assessment Inventory (PAI) são tendencialmente usadas como referenciais no âmbito da avaliação da validade de sintomas, uma vez que incluem escalas e índices que se destinam à avaliação específica da validade de relatos verbais (McGrath et al., 2010).

O Inventário Multifásico de Personalidade de Minnesota (MMPI-2; Butcher et al., 2001) é um instrumento de autorrelato composto por 567 itens que se destina à avaliação da personalidade. Trata-se de uma prova extensa, com elevado tempo de administração que requer treino na sua aplicação e cotação e, por isso, deverá ser administrada por examinadores experientes. Este instrumento é considerado o instrumento pioneiro para o início do estudo dos estilos de resposta, sendo também a medida mais extensivamente estudada

com este propósito (Barber-Rioja, Zottoli, Kucharski & Duncan, 2016; Faust, 1995; Wall et al., 2014; Wygant et al., 2018). São várias as escalas que procuram avaliar o domínio da validade. A escala F (Respostas Pouco Frequentes) foi a primeira das escalas a ser considerada neste sentido e pretende avaliar a tendência para o indivíduo se queixar de atitudes ou comportamentos pouco habituais na sequência de severa patologia, procurando colocar-se numa perspetiva prejudicial. Para além desta, o MMPI-2 inclui outras duas escalas de infrequência – a escala Fb, que procura analisar sintomas pouco frequentes de serem observados em indivíduos saudáveis, e a escala Fp, que procura examinar a existência de sintomas raramente encontrados em indivíduos com psicopatologia. A subescala mais investigada no âmbito da validade de sintomas é a FBS (Fake Bad Scale), evidenciando boas capacidades discriminativas em contextos clínicos e forenses (Heilbronner et al., 2009). A escala L (Mentira) procura avaliar tentativas do sujeito se colocar numa perspetiva moral e culturalmente favorável, negando ou escondendo imperfeições morais. A escala K (Natureza defensiva das respostas) visa a avaliação da tendência a controlar/limitar a comunicação de desconforto ou mal-estar, no sentido de comunicar uma perspetiva mais favorável de si mesmo, podendo evidenciar uma eventual resistência à avaliação. Por último, a escala ? (Interrogação), que procura verificar a existência de itens não respondidos ou duplamente assinalados. O MMPI-2 inclui, comparativamente à sua versão original, duas escalas que se destinam à avaliação da inconsistência das respostas: a escala TRIN, que avalia a tendência a responder em concordância (aquiescência), e a escala VRIN, que avalia a tendência de respostas ao acaso. Estas duas escalas são sobretudo utilizadas no sentido de detetar eventuais preferências por respostas Verdadeiro ou Falso. Contudo, importa referir que nenhuma destas escalas poderá, só por si, servir como referencial para o diagnóstico ou exclusão da possibilidade de simulação, dado que não avaliam a motivação subjacente à tendência de resposta (Nichols, 2011; Wygant et al., 2018). No caso particular das amostras forenses, o MMPI-2 revela um grande potencial informativo ao nível da avaliação da validade de sintomas (Barber-Rioja et al., 2016; Boccaccini, Murrie, & Duncan, 2006).

O Inventário de Avaliação da Personalidade (PAI; Morey, 1991, 2007) é uma outra medida de autorrelato que se destina à avaliação da personalidade e da psicopatologia, cuja aplicabilidade tanto em contextos clínicos como forenses tem sido reconhecida (Kucharski et al., 2007). É um instrumento composto por 344 itens, com formato de resposta em escala tipo Likert, que inclui quatro escalas de validade para avaliação da simulação/exagero de sintomas. A escala ICN (Inconsistência) atende a fatores como a pouca concentração e atenção na resposta ao inventário, a escala INF (Infrequência) analisa resposta pouco frequentes e/ou aleatórias, a escala NIM (Impressões Negativas) avalia relatos exagerados, sendo composta por itens que traduzem sintomas raramente presentes em psicopatologia real, e a PIM (Impressões Positivas), cujo propósito é a avaliação da tendência de resposta afirmativa (Kucharski et al., 2007; Morey & Lowmaster, 2010). Estas escalas oferecem informações a três níveis:

comunicação/distorção aleatória, distorção positiva (defensividade), e distorção negativa (falsificação) (Boccaccini & Hart, 2018). A aplicabilidade destas escalas, sobretudo da escala NIM, tem vindo a ser significativamente reconhecida no âmbito da avaliação da simulação e exagero de sintomas (Morey & Lowmaster, 2010), particularmente em contextos forenses (Kucharski et al., 2007). Para além destas escalas, o PAI inclui vários índices que procuram complementar a interpretação dos resultados. O MAL (Índice de Simulação) é baseado em oito configurações/características diferentes e procura avaliar de forma mais exaustiva os padrões de simulação, comparativamente à escala NIM. O RDF (Função Discriminativa de Rogers; Rogers, Sewell, Morey, & Ustad, 1996), por sua vez, corresponde à interpretação de uma combinação de itens que visa especificamente a deteção da simulação. Boccaccini e colaboradores (2006) procuraram avaliar o relato exagerado de sintomas numa amostra forense composta por indivíduos que seriam presentes a julgamento. Utilizando o PAI, os investigadores concluíram que a escala NIM era tão eficaz como o MMPI-2 na deteção de simuladores previamente detetados como tal pela SIRS. As conclusões de Kucharski e colaboradores (2007) seguem no mesmo sentido. Através do *design* de grupos conhecidos e utilizando a SIRS como critério externo, os investigadores concluíram pela aplicabilidade da escala NIM na deteção de sujeitos previamente assinalados como simuladores.

A utilidade do MMPI-2 e do PAI no âmbito da avaliação da validade de sintomas tem vindo a ser sustentada ao longo dos anos. No entanto, importa realçar que estes inventários não são sensíveis a simuladores sofisticados ou ao *coaching*, havendo uma significativa diminuição da sua capacidade de deteção quando os sujeitos estudam a composição das escalas ou as patologias, ou são instruídos a responder num dado sentido (Boccaccini et al., 2006; Rogers et al., 1993). Apesar de ser possível inferir conclusões no âmbito da validade de respostas a partir de escalas mais específicas presentes nos instrumentos clássicos atrás citados, esse não é o propósito das medidas, pelo que a informação que é possível retirar é mais limitada e menos objetiva na confirmação da credibilidade dos relatos (Simões, 2012; Van Dyke, 2013).

No sentido de colmatar estas fragilidades foram desenvolvidas medidas específicas para a avaliação deste domínio, baseadas em critérios psicométricos capazes de discriminar entre sujeitos com défices genuínos e aqueles que fingem – testes de validade de desempenhos (TVD) e testes de validade de sintomas (TVS) (Denney, 2007; Larrabee, 2007).

Contudo, devem ser considerados alguns pontos quanto à utilização e interpretação destas medidas específicas. Pontuações inválidas em apenas um destes instrumentos não deverão conduzir à invalidação direta do protocolo, devendo ser exploradas todas as possíveis causas para as classificações observadas – fadiga, falta de concentração, dificuldades na compreensão do instrumento. Para que seja questionada a validade do protocolo, a investigação sugere pontuações inválidas em duas ou três medidas de avaliação da validade, sendo, portanto, recomendada a inclusão de vários TVD/TVS no protocolo avaliativo (Larrabee, 2007; Slick et al., 1999; van Helvoort et al., 2019). O uso exclusivo destes instrumentos como forma de avaliação da

simulação reúne também algumas críticas. Apesar de a maioria das medidas construídas com este objetivo revelar bons índices de validade e fiabilidade, o seu uso exclusivo nos protocolos de avaliação poderá produzir um rácio de falsos positivos de 5% a 10% (Berry & Nelson, 2010). Para além disto, estes testes apenas captam o relato exagerado e pouco credível de padrões sintomáticos, não fornecendo quaisquer informações quanto à motivação por detrás do relato simulado (Merten & Merckelbach, 2013; van Impelen, Merckelbach, Jelicic, & Merten, 2014; van Impelen et al., 2017), pelo que é fundamental que a informação proveniente dos testes seja correlacionada com aquela que se retira dos restantes testes psicométricos, das entrevistas, dos dados processuais e do juízo clínico do profissional (Heilbronner et al., 2009).

A avaliação da validade de sintomas começa muito antes da administração de qualquer instrumento psicométrico através, por exemplo, da observação comportamental ou da entrevista clínica (Carr, 2015). A observação do comportamento revela-se fundamental na avaliação de sujeitos pouco comunicativos ou cooperantes. A entrevista clínica é essencial na interpretação e classificação da simulação, porém, isoladamente, não é suficiente para que se diagnostique o problema, dada a sua natureza subjetiva (Reid, 2000; citado em Carr, 2015). O juízo clínico profissional também não deve ser descurado, uma vez que a análise de eventuais incongruências ou contradições no perfil do sujeito poderão ser imperativas neste âmbito avaliativo (Rogers, 2018b). Para neutralizar estas dificuldades, a investigação aconselha o uso de múltiplas medidas e estratégias para a deteção da simulação (Larrabee, 2008; Lynch, 2004).

Assim, no sentido de standardizar o estudo da validade de sintomas foi necessária a exploração de técnicas que permitissem a deteção eficaz de perfis simuladores. Algumas destas estratégias têm vindo a ser incluídas nos instrumentos de avaliação destinados especificamente à avaliação da validade dos protocolos, auxiliando também o clínico na análise dos padrões de resposta aos instrumentos clássicos. No entanto, para que esta análise se dê de forma rigorosa, há que refletir sobre alguns pontos fundamentais. Em primeiro lugar, importa considerar que os estilos de resposta são tendencialmente encontrados em três domínios principais: doença mental, capacidades cognitivas, e sintomatologia médica (Rogers et al., 1993; Rogers, 2018b). Um indivíduo que procure simular fragilidades cognitivas optará por estratégias diferentes daquelas escolhidas por um indivíduo que pretenda simular a presença de uma patologia da personalidade. Tendo em conta estas diferenças, o examinador deverá considerar diferentes estratégias de deteção específicas para cada um destes domínios. No que concerne particularmente ao caso da simulação ou exagero de sintomas, têm sido apresentadas um conjunto de estratégias tanto no âmbito da simulação de sintomatologia e doença mental, como no âmbito da simulação de fragilidades (neuro)cognitivas.

Na seguinte tabela (Tabela 1) encontram-se sintetizadas um conjunto de estratégias que procuram facilitar a deteção de padrões de simulação (Berry & Nelson, 2010; Mittenberg et al., 2002; Rogers, 1990, 2018b).

Tabela 1: Estratégias para detecção da simulação

Sintomatologia/ Doença Mental	Sintomas Raros	Seleção de sintomas raramente presentes em indivíduos com psicopatologia real;
	Combinação de Sintomas	Seleção de sintomas que podem realmente ocorrer nas patologias, mas cuja presença em simultâneo é pouco provável;
	Sintomas Indiscriminados	Seleção de um grande número de sintomas (físicos e psíquicos) pertencentes a um vasto leque de patologias;
	Sintomas Severos	Seleção de um grande número de sintomas com extrema intensidade, frequência e duração, cuja severidade não é tendencialmente encontrada em psicopatologia real;
	Sintomas observados vs. Sintomas reportados	Comparação do padrão de respostas transmitidas pelo sujeito, com os dados provenientes da observação direta do comportamento, sendo analisadas as discrepâncias encontradas;
	Estereótipos erróneos	Baseada no facto de que existem diversos estereótipos associados às diferentes doenças mentais, sem base científica.
Fragilidades Cognitivas	Magnitude de erro	Sujeitos simuladores tendem a cometer mais erros do que aqueles já expectáveis de serem cometidos por sujeitos com reais dificuldades cognitivas.
	Performance atípica	Comparação do padrão de respostas entre instrumentos que procuram medir as mesmas capacidades.
	Curva de performance	Sujeitos simuladores podem apresentar mais erros em itens mais fáceis, acertando itens mais difíceis, ao passo que pacientes com patologia tendem a errar mais itens com o aumento da dificuldade.
	Violação das normas de instrução	Violação evidente das normas de instrução básicas de resposta ao instrumento.
	Efeito de teto	Falhar questões demasiado óbvias/fáceis, que geralmente são acertadas até por pessoas com défices cognitivos evidentes.

A seleção de sintomas raros e a combinação improvável de sintomas são as duas estratégias mais utilizadas no desenvolvimento dos TVS (Barber-Rioja et al., 2016; Rogers, 2018b). Apesar de ambas serem relativamente resistentes a possíveis simuladores treinados ou sofisticados, a combinação de sintomas revela-se mais resistente ao *coaching*, dado que os indivíduos podem reconhecer a sintomatologia associada a uma dada doença, mas não conhecer profundamente os padrões em que se manifesta. Esta “ingenuidade” no conhecimento da sintomatologia torna estes indivíduos mais facilmente identificáveis (Rogers, 2018b). A análise dos sintomas reportados pelo sujeito, em comparação com os sintomas observados, apenas deverá ser feita se o examinador for suficientemente treinado e conhecedor das patologias com detalhe, sob pena de se retirarem inferências erradas. Inconsistências ao nível da observação do comportamento, da postura e verbalizações ao longo das sessões de avaliação poderão fornecer informações importantes acerca da

intenção do sujeito em simular ou exagerar sintomas (Mittenberg et al., 2002). Uma evidente falta de cooperação e resistência perante as tarefas de avaliação, pautada por várias respostas do tipo “não sei” e desistências das tarefas nos itens mais fáceis, poderão ser também indicadores de que o sujeito não se estará a esforçar idealmente perante o processo avaliativo (Rogers, 2018b). No que toca à presença de estereótipos baseados no senso comum, importa considerar que mesmo os examinadores mais treinados possuem pré-conceitos relativos à expressão de determinada patologia, pelo que esta estratégia deverá ser apreciada com alguma cautela (Rogers, 2018b).

No que concerne à simulação de dificuldades intelectuais/cognitivas, as estratégias apontadas focam-se na análise dos padrões de resposta (desempenho) a instrumentos, e requerem a atenção do examinador para a postura (motivação, implicação ou desinteresse) do sujeito na realização destes testes (Simões, 2005; Conroy & Kwartner, 2006). Pacientes genuínos tendem a cometer alguns erros que serão, à partida, expectáveis, contudo, sujeitos simuladores tendem a errar itens pouco prováveis de serem errados por sujeitos com real patologia associada, pelo que a magnitude dos erros deverá ser analisada. Para além disto, há que considerar a curva de desempenho (*performance*), uma vez que se espera que sujeitos com ou sem défices cognitivos acertem itens mais fáceis e errem itens mais difíceis, sendo que um perfil inverso neste sentido poderá ser indicador de simulação (Simões, 2005). Goldstein (1945; citado em Rogers et al., 1993) postula que os simuladores não irão considerar a dificuldade das questões na seleção dos itens que escolhem errar, pelo que a comparação da curva de desempenho entre estes sujeitos e indivíduos que respondem honestamente, se revela eficaz. Uma das estratégias mais utilizadas relaciona-se com o efeito de teto, que se traduz na premissa segundo a qual sujeitos que simulam défices (neuro)cognitivos tendem a errar questões muito elementares de informação (como “O que é maior? Um rato ou um cavalo?”) com o objetivo de realçar determinadas dificuldades. No entanto, há que considerar que indivíduos com graves lesões cerebrais poderão efetivamente errar determinadas questões sem que tal corresponda à intenção de simular. Para além da análise do perfil de respostas intra-teste, importa também analisar o desempenho entre diferentes testes. É comum que num perfil de simulação se evidenciem discrepâncias entre provas que avaliam os mesmos processos mentais (memória, atenção, ...), todavia, importa considerar que pacientes com reais lesões cerebrais podem evidenciar maiores dificuldades num teste específico, respondendo mais facilmente a um outro instrumento que avalie o mesmo domínio, pelo que esta estratégia carece ainda de estudos empíricos suficientemente validados (Rogers et al., 1993; Rogers, 2018b).

O campo mnésico é uma das áreas de avaliação cognitivas em que tendencialmente se encontram mais queixas simuladas ou exageradas (Russeler, Brett, Klaue, Sailer, & Munte, 2008). Por essa razão, é fundamental que os profissionais tenham em consideração as estratégias acima mencionadas e alguns pontos específicos deste domínio. Os examinadores deverão estar atentos a eventuais perfis de resultados que não façam sentido sob o ponto de vista (neuro)psicológico nomeadamente: défices muito

acentuados em tarefas de evocação diferida, comparativamente à evocação imediata; as pontuações nas provas de evocação serem marcadamente superiores aos resultados das provas de reconhecimento, uma vez que nos problemas mnésicos reais esta relação tende a ser inversa (Hanley, Baker & Ledson, 1999; Simões, 2005).

Para além das áreas acima consideradas, o clínico deverá estar também ciente de outras áreas em que os sujeitos tendem a exagerar mais sintomatologia. Destas áreas, destacam-se: dores, vertigens e tonturas; depressão, nervosismo e irritação; perturbações mnésicas; confusão geral e falta de compreensão; distração durante o processo avaliativo; dislexia; alterações de personalidade; cegueira ou perda parcial de visão; torpor; mobilidade limitada ou reduzida amplitude de movimentos (Iverson & Binder, 2000).

No entanto, é possível que mesmo utilizando os indicadores mais adequados, alguns sujeitos consigam produzir ou exagerar sintomatologia, sem que tal seja detetado pelos instrumentos psicométricos ou pelo juízo clínico do examinador. Sullivan e King (2010) apontam para que o formato de resposta de alguns instrumentos, como as *checklists* de tipo Likert ou a resposta dicotómica, possam levar a que os respondentes assinalem sintomas que não reportariam em questões, por exemplo, de resposta aberta. Para além disto, os indivíduos tendem a expressar a sintomatologia da doença tendo em conta a forma como a percebem (Pollock, et al., 1997). Como tal, se o sujeito tiver um conhecimento profundo acerca de determinada patologia, a forma como reproduzirá sintomas será muito mais sofisticada e não tão perceptível pelos instrumentos. O nível cultural de conhecimento das psicopatologias e o acesso a meios científicos de diagnóstico formal poderá contribuir para o aperfeiçoamento da técnica de simulação – simuladores sofisticados. De forma a contrariar esta possibilidade, os examinadores terão de conhecer as patologias e as suas manifestações com alguma profundidade (Rogers, 2018b).

Estas estratégias poderão ser um bom ponto de partida para uma avaliação mais exaustiva da simulação ou exagero de sintomas, uma vez que se constituem como pistas importantes para o examinador. Contudo, algumas variáveis contextuais, estando presentes, deverão contribuir para que a atenção do clínico seja reforçada perante a possibilidade de simulação ou exagero de sintomas, nomeadamente: a presença de incentivos externos (como a obtenção de baixa médica, reforma antecipada, atenuação da pena,...); historial irregular de emprego – lesões incapacitantes, ausências prolongadas; comportamentos de desonestidade enraizados; profundo conhecimento dos contornos legais relevantes para o seu caso; historial de exposição a processos litigiosos; experiências prévias com situações de prestação de cuidados de saúde ou companhias de seguro (Drogin & Williams, 2018, Simões, 2005).

Concluindo, no sentido de neutralizar esta tendência, a deteção da simulação deve ser multi-método, recorrendo a vários indicadores: vários tipos de provas psicométricas, entrevistas, e estratégias de identificação múltipla (Bush et al., 2005; Heilbronner et al., 2009; Mittenberg et al., 2002; Rogers, 2018).

2.1 Testes de Validade de Desempenho (TVD)

Os TVD são medidas direcionadas para a avaliação do desempenho ou do esforço em testes (neuro)psicológicos, i.e., são instrumentos cujo propósito é avaliar se o sujeito procura demonstrar as suas verdadeiras capacidades (Larrabee, 2014; Van Dyke, Millis, Axelrod & Hanks, 2013) e, portanto, se as pontuações obtidas nos testes de avaliação se traduzem numa medida rigorosa das suas reais aptidões (Simões et al., 2017). Os TVD baseiam-se no pressuposto segundo o qual desempenhos inferiores à probabilidade de resposta ao acaso num grau estatisticamente significativo, ou marcadamente abaixo dos pontos de corte definidos para o teste, são indiciadores de respostas enviesadas – efeito de teto (Bender & Frederick, 2018). Assim, um perfil de resultados abaixo deste parâmetro será indicador de uma intenção de não responder corretamente (Boskovic, Merckelbach, Merten, Hope & Jelacic, 2019; Merten & Merckelbach, 2013).

Dois dos TVD mais empiricamente investigados e validados são o Test of Memory Malinger (TOMM; Tombaugh, 1996) e o Rey 15 – Item Memory Test (FIT; Rey, 1964; Boone et al., 2002)

O TOMM (Tombaugh, 1996) é o teste mais usado no âmbito da avaliação da validade de desempenho (Martin et al., 2019). Trata-se de uma medida de escolha forçada, constituída por uma tarefa de reconhecimento visual que visa diferenciar os desempenhos de sujeitos com défices cognitivos (mnésicos) reais daqueles que procuram alterar o seu desempenho (Tombaugh, 1996; Heubrock & Petermann, 1998; Simões et al., 2017). O TOMM tem vindo a ser reconhecido como um instrumento válido neste âmbito, uma vez que não é facilmente reconhecido pelos sujeitos como um TVD devido ao elevado número de estímulos que assegura a sua validade facial enquanto teste de aprendizagem e memória (Simões et al., 2017). Em Portugal, os estudos têm validado a utilização deste instrumento em várias amostras, nomeadamente crianças, adultos, idosos e em diversas de amostras forenses. No que concerne particularmente a estudos realizados com amostras prisionais, as conclusões validam a inclusão do TOMM nos protocolos de avaliação (Simões et al., 2017). No entanto, algumas limitações apontadas a este instrumento sugerem que os seus resultados apenas permitem concluir a presença ou não de falsificação/exagero de sintomas, sem que seja possível interpretar a motivação por detrás desse relato. Logo, sendo este um critério central, não é possível diagnosticar um perfil de simulação com base exclusivamente nos resultados desta medida (Simões et al., 2017).

O Rey 15-IMT (FIT; Rey, 1964; Boone et al., 2002) é um teste de escolha não forçada, desenvolvido com o objetivo específico de avaliar as queixas cognitivas e o desempenho associados à simulação de sintomas e ao esforço insuficiente (Simões, 2012; Vilar et al., 2017). À semelhança do TOMM, este instrumento possui uma excelente validade facial, assumindo-se como uma medida relevante a incluir nos protocolos de avaliação tanto clínicos, como forenses (Boone et al., 2002; Vilar et al., 2017). Particularmente no que toca ao contexto prisional, um estudo realizado, em Portugal, com uma amostra de reclusos conclui que esta medida apresenta uma maior sensibilidade ao efeito da idade e escolaridade, pelo que o TOMM

se assume como o instrumento com maior potencial a ser empregue neste meio (Soares, 2013; Vilar et al., 2017).

2.2 Testes de Validade de Sintomas (TVS)

Os TVS procuram avaliar se o relato de sintomas do sujeito corresponderá à sua vivência e capacidades reais (Greve et al., 2008; Simões & Sousa, 2008; Van Dyke et al., 2013). Estes instrumentos são compostos por itens que procuram refletir sintomas bizarros ou raramente encontrados em indivíduos com verdadeira psicopatologia (Bender & Frederick, 2018), uma vez que se baseiam no pressuposto de que sujeitos que pretendem simular irão reportar excessivamente sintomas, mesmo que de forma extrema (Boskovic et al., 2019). Mais recentemente, a designação de TVS tem vindo a ser aplicada apenas a testes que se destinam a avaliar a falsificação de sintomatologia/doença psiquiátrica, no sentido de os distinguir dos instrumentos que procuram avaliar a produção ou exagero de défices (neuro)cognitivos (TVD). Esta distinção tem-se revelado importante, dado que o exagero de sintomatologia e o esforço reduzido nas provas cognitivas nem sempre se correlacionam (Bender & Frederick, 2018). A maioria destes instrumentos baseiam-se no autorrelato como medida de avaliação (Heilbronner et al., 2009). Apesar de existirem diversos instrumentos elaborados neste propósito, apenas alguns são comumente utilizados e aceites pela generalidade da comunidade científica (Greve et al., 2008). Os TVS mais referidos na literatura são a Structured Interview of Reported Symptoms (SIRS; Rogers, Bagby, & Dickens, 1992; Amaral, 2007; SIRS-2; Rogers Sewell, & Gillard, 2010), o Structured Inventory of Malingered Symptomatology (SIMS; Smith & Burger, 1997; Widows & Smith, 2005; versão portuguesa de Simões et al., 2017), e, mais recentemente, o Self-Report Symptom Inventory (SRSI; Merten et al., 2016.). Em Portugal deve igualmente ser referenciada a Escala de Validade de Sintomas (EVS-2, Cunha, 2011; Simões, Almiro, Nunes et al., 2017).

A SIRS (Rogers, Bagby, & Dickens, 1992; Amaral, 2007) é uma entrevista estruturada composta por 172 itens que visa a avaliação da simulação de sintomas. É composta por oito escalas primárias, cada uma representativa de uma estratégia diferente que os sujeitos podem assumir na simulação de uma psicopatologia, e cinco escalas suplementares (Rogers, 1990; Simões, 2012). A SIRS encontra-se validada tanto em *designs* de simulação análogos como em *designs* com grupos conhecidos, evidenciando em ambos bons valores de consistência interna e fiabilidade (Barber-Rioja et al., 2016). Trata-se de uma medida válida e fidedigna a ser usada em vários contextos, nomeadamente clínicos, forenses, e prisionais, reunindo bastante consenso na sua utilização (Edens et al., 2007; Rogers, Gillis, Dickens & Bagby, 1991; Sellbom et al., 2010). Por se tratar de uma entrevista relativamente extensa, que requer profissionais suficientemente treinados e algumas horas de aplicação poderá não ser a primeira escolha a incluir nos protocolos de avaliação (Clegg, Fremouw & Mogge, 2009). No entanto, o formato de entrevista confere também algumas vantagens ao instrumento. Isto

porque é possível que alguns resultados sejam enviesados devido ao formato de resposta dicotômica, uma vez que poderá haver uma maior tendência em responder de forma aleatória ou impulsiva (Lilienfeld et al., 2013). Neste sentido, caso a análise dos padrões de resposta aos restantes instrumentos suscite questões quanto à validade dos relatos, a SIRS constitui-se como uma excelente medida para avaliar mais profundamente este domínio (Boccaccini et al., 2006; Edens et al., 2007; McDermott & Sokolov, 2009).

A EVS-2 (Simões et al., 2017) foi desenvolvida em Portugal com o objetivo de dar resposta à crescente necessidade de recurso a vários instrumentos para avaliação da simulação de sintomas (Simões, 2012). Trata-se de uma medida cujo objetivo principal é analisar a simulação ou o exagero de sintomas psicopatológicos e neurocognitivos, tendo sido desenvolvida com a intenção de ser utilizada como complemento ao SIMS e à SIRS (Simões, et al., 2017). A EVS-2 é um instrumento breve de autorrelato, com formato de resposta dicotômico – Verdadeiro/Falso, constituída por 48 itens, distribuídos por 3 domínios – Psicose (P), Perturbações Emocionais (PE), Problemas Cognitivos (PC) (Ferreira, 2017). Apesar de ter sido construída com base nos mesmos construtos teóricos que o SIMS, os instrumentos não têm nenhum item em comum sendo, portanto, recomendada a sua utilização conjunta, de forma complementar (Simões, Almiro, Nunes et al., 2017).

O SIMS (Widows & Smith, 2005) é um instrumento de autorrelato com formato de resposta dicotômica – Verdadeiro/Falso, constituído por 75 itens e 5 subescalas que avaliam os seguintes domínios: Psicose (P), Défice Neurológico (NI), Perturbações Mnésicas (AM), Capacidade Intelectual Reduzida (LI), e Perturbações Afetivas (AF) (Smith & Burger, 1997). Para além de ser o instrumento de autorrelato mais utilizado a nível mundial para a avaliação da simulação (De Marchi & Balboni, 2018; Vitacco, Rogers, Gabel, & Munizza, 2007), é também o mais usado em Portugal e dispõe de vários estudos de validação desenvolvidos em contexto forense (cf. Simões, Almiro, Mota et al., 2017). Este instrumento foi construído com base no pressuposto de que os sujeitos simuladores pontuam mais em itens que transpareçam sintomas bizarros ou atípicos, em comparação a sujeitos com efetivas queixas psicopatológicas (Seron, 2014).

O SIMS tem sido objeto de estudo em Portugal e os resultados são genericamente sugestivos da sua utilidade (Simões, 2012). Trata-se de uma medida que abrange uma grande variedade de psicopatologia e apresenta uma excelente consistência interna em todas as subescalas e no resultado total (Vitacco et al., 2006; Simões, Almiro, Mota et al., 2017). Para além disto, o SIMS encontra-se validado tanto em investigações que utilizam o *design* de simulação análogo, como o *design* de grupos conhecidos, sendo um dos poucos inventários com resultados válidos neste último *design* (Vitacco et al., 2007; Vitacco & Rogers, 2005). No entanto, a investigação recente tem sido bastante crítica no que toca à utilidade deste instrumento nos vários contextos. Inicialmente, o SIMS foi pensado como um instrumento de triagem para incluir em avaliações forenses (Merckelbach & Smith, 2003; Smith & Burger, 1997), o que justifica a inclusão apenas de sintomas ou combinações de sintomas atípicas/bizarros e não de sintomas genuínos, encontrados em

indivíduos com psicopatologia real (Giger & Merten, 2019). Esta configuração da escala não permite que défices cognitivos menos severos e menos incapacitantes sejam detetados. Tal constitui-se como uma significativa limitação do instrumento uma vez que não permite a deteção de lesões/perturbações moderadas ou leves (Merten, et al., 2016; van Impelen et al., 2014). O estudo de Cernovsky, e colaboradores (2019) conclui que a maioria dos itens que compõem o SIMS não permitem que se faça uma diferenciação rigorosa entre sujeitos simuladores e não simuladores. Um outro estudo conduzido pelos mesmos autores, avaliou uma amostra de 98 sobreviventes de acidentes de carro, e as conclusões sugerem que mais de metade dos sintomas representados na escala podem efetivamente estar presentes nestes sujeitos, sem que estes tenham qualquer intenção de simular ou exagerar. Para além disto, praticamente nenhum dos itens foi eficazmente capaz de diferenciar sintomas simulados e sintomas genuínos manifestados pelos indivíduos (Cernovsky, Ferrari & Mendonça, 2019). Por último, importa referir que, comparativamente com as restantes medidas de autorrelato utilizadas na avaliação da validade de sintomas, o SIMS é mais vulnerável ao *coaching* (Hunt, 2013), uma das maiores ameaças à validade e credibilidade das avaliações (Springman & Vanderberg, 2009).

Estas limitações tornam o SIMS mais facilmente reconhecível como TVS (Bonkovic et al., 2019), não se assumindo como uma medida eficaz na deteção de simuladores sofisticados (De Marchi & Balboni, 2018; Geurten et al., 2018; Merten, et al., 2016; van Helvoort et al., 2019).

Mais recentemente, e no sentido de dar resposta às fragilidades apontadas ao SIMS e aos restantes instrumentos de exame da validade de sintomas (van Helvoort et al., 2019), foi desenvolvido um novo instrumento. O SRSI (Merten et al., 2016) é um questionário de autorrelato com formato de resposta dicotómica (verdadeiro/falso), que combina 5 escalas de Sintomas Genuínos com 5 escalas de Pseudosintomas. As primeiras correspondem a sintomas habitualmente relatados em indivíduos com real patologia, ao passo que as segundas correspondem a sintomas estranhos, atípicos ou raramente encontrados em psicopatologia real. As escalas de Sintomas Genuínos avaliam vários domínios que integram diferentes subescalas, como a cognição, depressão, dor, dimensão somática e ansiosa. As escalas de Pseudosintomas incluem subescalas que se destinam à avaliação do domínio cognitivo, motor, sensorial, ansioso/depressivo, e da dor (Merten et al., 2016). De acordo com os autores, é expectável que pacientes honestos, com genuína sintomatologia, pontuem mais nas escalas de Sintomas Genuínos e menos nas escalas de Pseudosintomas, contrariamente ao que se espera que aconteça com sujeitos simuladores (Boskovic et al., 2019; Geurten et al., 2018; Jelacic, Hessels, & Merckelbach, 2006; Merten et al., 2016;). O SRSI foi construído com base em algumas estratégias de deteção de simulação de sintomatologia/doença mental, especificamente a seleção de sintomas raros, improváveis e a combinação de sintomas. (Dwarkadas, 2018; Merten et al., 2016; Van Helvoort et al., 2019).

Num total de 107 itens, 100 compõem as escalas e 7 destinam-se à avaliação da atitude perante o teste: 2 itens iniciais, com o objetivo de

determinar se o indivíduo investirá nas respostas ao instrumento, e 5 outros itens, uniformemente distribuídos, com o objetivo de detetar respostas ao acaso, particularmente a tendência para a resposta afirmativa.

O SRSI tem vindo a ser explorado em vários contextos e em diferentes países. Os resultados sugerem que tanto as subescalas de Sintomas Genuínos como as de Pseudosintomas são eficazes na diferenciação de sujeitos simuladores e não simuladores (Geurten et al., 2018; Giger & Merten, 2018). No âmbito forense, van Helvoort e colaboradores (2019), ao administrar o instrumento a uma amostra composta por 40 reclusos internados num hospital psiquiátrico, concluíram que em comparação ao SIMS, a escala de Pseudosintomas do SRSI parece relativamente imune à presença de psicopatologia, não assumindo como simulação os casos de real psicopatologia. No contexto português foram já concluídas algumas investigações, tanto em contexto prisional, como com amostras comunitárias (Domingues, 2019; Dwarkadas, 2018; Pinheiro, 2019). Os resultados seguem no mesmo sentido, sustentando o SRSI como instrumento potencialmente útil a incluir nos protocolos de avaliação. Tanto para a amostra comunitária, como para a amostra prisional, os estudos indicam como relações positivas e significativas entre a Escala de Pseudosintomas e o resultado total do SIMS, o que seria já expectável considerando que medem os mesmos construtos (Dwarkadas, 2018; Pinheiro, 2019).

No âmbito da avaliação de sintomas, o SRSI é, claramente, um instrumento promissor a incluir em protocolos de avaliação forenses (Geurten et al., 2018; Giger & Merten, 2019; Merten et al., 2016; Pinheiro, 2019; Van Helvoort et al., 2019), apesar de serem ainda necessários mais estudos que corroborem a utilidade do instrumento em contexto clínico (Boskovic et al., 2019). Assim a continuação da realização de estudos com o SRSI em outras populações e utilizando diferentes *designs* é essencial, dado que poderá viabilizar a sua aplicabilidade enquanto TVS eficaz.

Finalmente, importa ter em consideração algumas limitações dos TVS. Estes instrumentos exigem um mínimo de capacidades cognitivas preservadas, ao nível da leitura e compreensão, pelo que não poderão ser administrados a indivíduos com défices intelectuais severos (Dandachi-FitzGerald et al., 2020). Outra preocupação associada ao uso destes testes baseia-se no facto das medidas de autorrelato não avaliarem respostas aleatórias ou inconsistentes, nem medirem a desejabilidade social (Merten et al., 2016). Estas fragilidades devem ser tidas em conta e meticulosamente analisadas, uma vez que os resultados dos TVS devem ser informativos, independentemente do contexto de avaliação (Conroy & Kwartner, 2006; Merten et al., 2016).

III – *Designs* de validação: o *design* de grupos conhecidos

Ao longo das últimas décadas têm surgido algumas preocupações quanto aos *designs* usados na validação e estudo de alguns instrumentos e comportamentos de simulação.

O *design* de simulação análogo é o mais utilizado, ao nível global, para o estudo da validade de sintomas (Rogers, 2018a). A maioria destes estudos implica a distribuição aleatória de uma amostra não clínica em duas condições: uma subamostra é instruída a responder de forma honesta, e a outra é incitada a responder segundo um guião (*script*) associado a determinado estilo de respostas, com o objetivo de comunicar sintomas identificados com a definição de problema ou doença particulares. No caso específico da simulação, é pedido aos sujeitos que finjam determinados défices ou queixas. As médias dos resultados obtidos nas condições são então comparadas no sentido de perceber se são observadas as diferenças expectáveis entre as condições. As respostas são depois comparadas com as de uma amostra clínica. Se forem encontradas diferenças significativas entre as duas condições, assume-se que o instrumento em causa será capaz de diferenciar respostas honestas de respostas simuladas (Heilbronner et al., 2009; Rogers, 2018a). Este *design* assume potencialidades ao nível da validade interna, uma vez que os resultados observados apenas poderão ser atribuídos às medidas em análise, e não a outras variáveis que não foram tratadas. No entanto, uma vez que são trabalhados cenários específicos com um determinado grupo, este *design* revela fragilidades no âmbito da validade externa, uma vez que é difícil a generalização das conclusões a outros grupos e condições (Rogers, 2018c; Rogers & Cruise, 1998; Sellbom et al., 2010; Wisdom, Callahan & Shaw, 2010). Para além disto, a utilização deste *design* poderá sobrestimar a validade dos TVS, uma vez que pontuações elevadas nestes testes não implicam necessariamente a intencionalidade de simular, pelo que esta relação deverá ser cautelosamente analisada sob pena de aumentar exponencialmente o rácio de falsos positivos (Lilienfeld et al., 2013). Rogers (2018a) acrescenta que a maioria destes estudos se mostram pouco fidedignos dada a exclusiva utilização de populações não clínicas, sobretudo amostras compostas por estudantes que, apesar de serem as mais oportunas, não permitem comparações com amostras com efetiva psicopatologia.

O *design* de grupos conhecidos apresenta-se como uma potencial alternativa, sendo mesmo considerado por muitos autores como preferencial no estudo e medição dos níveis de simulação (Rogers, 2018a). Este método começou por ser explorado por Cronbach e Meehl (1955) no sentido de testar a validade de construto, e baseia-se no pressuposto segundo o qual um instrumento será válido quando discrimina eficazmente um grupo de indivíduos conhecidos por ter uma característica específica de um grupo que não a possui. De uma forma geral, este *design* envolve a comparação de grupos presumíveis de se diferenciarem com base no constructo em análise (Hattie & Cooksey, 2015). A sua aplicação no âmbito da simulação de sintomas implica a análise de sujeitos identificados como simuladores, com base em critérios independentes daqueles a que o estudo se propõe. O perfil destes indivíduos é comparado com uma amostra de controlo, composta por

sujeitos que não se prevê que estejam a simular. Os critérios para a constituição dos grupos são geralmente estabelecidos por profissionais especializados na investigação da validade de sintomas, todavia, não havendo essa possibilidade a investigação recomenda a utilização de medidas de avaliação da simulação, suficientemente validadas e aferidas para a população em causa, acautelando a utilização de pontos de corte adequados (Rogers, 2018a; Rogers & Cruise, 1998). As recomendações vão no sentido de os grupos serem representativos da população em análise, devendo a sua constituição basear-se em critérios tidos como válidos e suficientemente aferidos para a amostra em causa e, por isso, com taxas de certeza iguais ou superiores a 90% (Hattie & Cooksey, 2015; Rogers, 2018a). Este *design* é composto por duas fases: em primeiro lugar, devem estabelecer-se grupos de critério (por exemplo, um grupo composto por sujeitos com efetiva psicopatologia, e outro grupo de sujeitos simuladores); e em segundo lugar, devem comparar-se esses grupos, ao nível de semelhanças e diferenças para que daí se possam inferir conclusões. A fase mais desafiante é indubitavelmente a primeira, a constituição dos grupos, uma vez que erros a este nível comprometerão toda a investigação (Rogers, 2018a).

Apesar de não ser um método tão explorado, algumas investigações têm vindo a aplicar o *design* de grupos conhecidos em vários contextos. No âmbito das investigações relacionadas com o MMPI-2, Bianchini e colaboradores (2008) procuraram estudar as escalas de validade deste instrumento na deteção da simulação em pacientes com dor crónica, e os resultados sugerem a eficácia destas escalas na diferenciação de sujeitos simuladores e não simuladores.

No âmbito da validade de desempenho, Greve e colaboradores (2006) aplicaram o TOMM a uma amostra composta por indivíduos que sofreram um traumatismo craniano, comparando estes sujeitos com uma amostra de sujeitos previamente classificados como simuladores. A utilidade deste instrumento foi também corroborada na sua aplicabilidade à população espanhola, através deste mesmo *design* (Vilar-López, Gómez-Río, Santiago-Ramajo, Rodríguez-Fernández, Puente & García, 2008).

Especificamente no que abrange o contexto forense, algumas investigações deverão ser mencionadas. Sellbom e colaboradores (2010) procuraram investigar a utilidade do MMPI-2-RF na deteção da simulação neste contexto. Utilizando a SIRS como critério externo, os investigadores aplicaram o inventário a 125 indivíduos, e os resultados são indicadores das potencialidades desta versão do MMPI-2 na avaliação da simulação ou exagero de sintomas. De Marchi e Balboni (2018) utilizaram o *design* de grupos conhecidos e o *design* de simulação análogo para explorar a eficácia do MMPI-2, do SIMS e da escala NIM na deteção da simulação numa amostra forense. Utilizando o *design* de grupos conhecidos, apenas o SIMS revelou distinguir significativamente sujeitos simuladores de não simuladores, contrariamente ao que se verificou no *design* de simulação análogo, em que todos os instrumentos revelaram resultados promissores. Considerando que a análise da curva ROC confirmou a maior sensibilidade do SIMS, poder-se-à concluir que o *design* de grupos conhecidos se revelou mais minucioso na classificação dos instrumentos.

As potencialidades deste *design* têm vindo a ser reconhecidas e sustentadas. Em comparação com o *design* análogo de simulação, o *design* de grupos conhecidos revela-se vantajoso no sentido em que a amostra utilizada encara problemas, incentivos e consequências reais (Rogers, 2018a). No entanto, a grande vantagem deste *design* prende-se, sobretudo, com o aumento da capacidade de generalização das conclusões (Wisdom et al., 2010). A investigação tem, de facto, corroborado o *design* de grupos conhecidos como aquele que mais contribui para a avaliação da validade de constructo, sendo que a validade externa é assim melhor avaliada comparativamente ao *design* de simulação análogo (Langevin, 2009; Rogers et al., 1993; Rogers & Cruise, 1998; Vitacco et al., 2006).

Contudo, nenhum método é isento de fragilidades. Poderão existir algumas fragilidades no âmbito da validade interna, que se constitui como relativamente fraca, uma vez que os investigadores não têm um grande controlo acerca do esforço e compromisso dos sujeitos perante o processo de avaliação. No entanto, a estandardização dos procedimentos e a utilização de critérios e instrumentos credíveis para a constituição dos grupos poderá diminuir este efeito (Rogers, 2018a; Rogers & Cruise, 1998; Slick et al., 1999; Wisdom et al., 2010). Para além de não ser um *design* extensivamente usado e, por isso, alvo de um número relativamente pequeno de investigações, os métodos sob os quais se baseiam as constituições dos grupos, são alvo de algumas recomendações por parte da comunidade científica (Hattie & Cooksey, 2015). Uma das limitações mais evidentes atribuída a este método prende-se com a utilização de apenas um instrumento como critério (De Marchi & Balboni, 2018). Rogers (2018a) analisa vários estudos com este *design* e as conclusões corroboram a premissa de que deverão ser utilizados mais do que um TVS na constituição dos grupos de forma a aumentar a confiança nos resultados. Para além disto, importa ainda referir que, idealmente, os instrumentos usados como critérios para a constituição dos grupos deveriam ser baseados tanto em parâmetros psicométricos como clínicos, de forma a garantir uma melhor avaliação dos perfis (De Marchi & Balboni, 2018).

IV – Avaliação da Simulação em Contexto Forense

Em contexto clínico é esperado que os indivíduos relatem sintomas de forma genuína e honesta, dada a relação terapêutica ser baseada na empatia e na confiança, e assumir finalidades terapêuticas. O objetivo do profissional de saúde é sempre o de minimizar o sofrimento do paciente, no sentido de promover a sua qualidade de vida e bem-estar, pelo que será benéfico e útil ao doente que comunique honestamente com o clínico. No entanto, em contexto forense, o cliente das avaliações raramente é o indivíduo avaliado, mas sim o Tribunal ou a instituição legal. Por essa razão, a sinceridade dos relatos poderá traduzir-se em consequências legais que não vão ao encontro dos objetivos do sujeito, pelo que a credibilidade dos sintomas reportados pode ser posta em causa (Faust, 1995; Giger & Merten, 2019). Segundo o estudo de Martin e colaboradores (2015) realizado com uma amostra de 316 (neuro)psicólogos, as perturbações psiquiátricas são tidas como a principal causa para a

invalidação dos protocolos de avaliação em contexto clínico, ao passo que esta atribuição causal incide na simulação ou relato exagerado de sintomas, em contextos forenses.

Archer e colaboradores (2006) conduziram uma investigação com o objetivo de perceber quais os instrumentos psicológicos tendencialmente usados pelos examinadores forenses. As conclusões sugerem que, no âmbito da avaliação da simulação, a SIRS seria a medida mais frequentemente incluída nos protocolos, com 30% dos profissionais inquiridos a incluir este instrumento. Contudo, 79% da amostra revela apenas incluir o MMPI no sentido da avaliação da validade de sintomas. Estes valores refletem a pouca importância que, há pouco mais de uma década, era dada à avaliação da credibilidade dos relatos. Assim, o pouco investimento e sofisticação usados na avaliação da simulação poderá levantar questões quanto ao facto de as prevalências encontradas, até à altura, estarem subestimadas. Contudo, na última década, a comunidade científica tem reunido consenso quanto à importância desta avaliação, tendo vindo a promover o desenvolvimento de medidas específicas (TVD/TVS). A importância de recorrer a vários instrumentos, suficientemente validados, é transversal a qualquer contexto avaliativo. Contudo, tratando-se o meio forense de um contexto cujas avaliações são particularmente alvo de escrutínio, é fundamental que os relatórios sejam fundamentados com rigor e base científica, devendo ser comunicadas informações provenientes de instrumentos de avaliação validados, sob pena de o perfil do sujeito ser inválido (Rogers, 2018b).

O âmbito forense abrange múltiplos contextos específicos que, apesar dos vários elementos comuns, diferem em certos aspetos. Assim, é expectável que, mudando o contexto, as motivações para produzir ou exagerar intencionalmente sintomatologia variem também. No contexto médico-legal, os relatos simulados poderão ter por base motivações relacionadas com a tentativa de evitar ir a julgamento, ou apresentar-se inocente por inimizabilidade dos factos. Na verdade, as prevalências sugerem que 8 (Cornell & Hawk, 1989) a 17.4% (Rogers et al, 1998) dos indivíduos que tentam evitar ir a julgamento, estarão a simular sintomatologia (McDermott et al., 2013). Em Direito Civil as prevalências são tendencialmente maiores, comparativamente ao contexto médico-legal (McDermott et al., 2013), sendo que a simulação de sintomas poderá ser motivada pela procura de indemnizações/compensações financeiras (Greve et al., 2008) ou determinados benefícios, em que o sujeito, alegando alguma incapacidade ou fragilidade, poderá alcançar a reforma antecipada, baixa médica, alterações das condições de trabalho ou acesso a fármacos específicos. Larrabee (2003) encontrou prevalências que atingem os 40% em sujeitos que procuravam compensação financeira alegando lesões cerebrais. Um estudo conduzido por Greve e colaboradores (2009) concluiu que, dos 508 pacientes que alegavam sofrer de dor crónica na busca por uma dada compensação financeira, 20 a 50% da amostra estaria a simular sintomatologia. Ainda neste sentido, aproximadamente 30% dos veteranos de guerra que alegavam sofrer de Perturbação de Stress Pós-Traumática foi indicada como estando a simular (Frueh, Hamner, Cahill, Gold, & Hamlin, 2000; citados em McDermott et al.,

2013). Por sua vez, em Direito de Família e Menores, a modificação dos relatos poderá ser motivada pela alteração ou manutenção das medidas que abrangem o exercício das responsabilidades parentais (Simões, 2012; Rogers, 2018a).

Particularmente no que concerne ao contexto penal e prisional, o recurso à simulação de sintomas poderá ser útil aos indivíduos na medida em que poderá contribuir para a justificação, desculpabilização ou atenuação das consequências do comportamento criminal, nomeadamente através da redução da pena aplicada, da ilibação/atenuação da culpa, ou obtenção de cuidados de saúde especiais (Denney, 2007; Resnick & Knoll, 2018; Rogers, 2018b; Wall et al., 2014). Um sujeito que se encontra em reclusão poderá simular psicopatologia visando receber melhores condições, como colocação em celas mais confortáveis ou alas menos restritivas, isenção de exercer atividade laboral (Wall et al., 2015) ou a prescrição de medicação (McDermott & Sokolov, 2009; Resnick & Knoll, 2018). Poderá dar-se também o caso de os reclusos não mostrarem evidente motivação para simular défices, exagerando sintomatologia apenas por falta de empenho ou aborrecimento face ao processo de avaliação (Denney, 2007), ou meramente por busca de atenção e divertimento na tentativa de manipulação do sistema de justiça (Wall et al., 2015; Walters, 2006), não sendo correto atribuir um diagnóstico de simulação nestes casos.

Das inúmeras particularidades da população reclusa, a investigação carece, ainda, de estudos suficientes que permitam caracterizar com rigor esta população quanto à simulação e exagero de sintomatologia. Contudo, alguns estudos permitem inferir importantes conclusões. Apesar de serem observadas taxas elevadas de simulação em reclusos já condenados, o maior risco de simulação associa-se a presos preventivos ou com processos penais pendentes (Denney, 2007). Isto porque, não havendo ainda uma conclusão do processo, é expectável que os indivíduos tentem atenuar ou mesmo evitar a pena de prisão efetiva, ou garantir melhores condições para o seu cumprimento (Rogers, 2018b). Norris e May (1998) concluíram que os reclusos simuladores seriam tendencialmente mais jovens e menos escolarizados, comparativamente aos reclusos respondentes honestos. Esta assunção poderá justificar-se tendo em conta que uma baixa escolaridade se traduzirá em manifestações menos sofisticadas das patologias ou défices, que serão, portanto, mais facilmente reconhecidas e detetadas pelos instrumentos. Para além disto, a impulsividade e a tendência para o risco, característicos da população mais jovem, poderão justificar o facto de estes serem mais facilmente reconhecíveis nas tentativas de simular sintomas. Os dados da investigação não são consensuais quanto à ligação entre a severidade dos crimes e a tendência para a simulação de sintomas. Taylor e Kopelman (1984; citados em McDermott et al., 2013) encontraram evidências de que o crime de homicídio poderá relacionar-se com a probabilidade de optar por um relato simulado. Contudo, McDermott e colaboradores (2013) não encontraram nenhuma associação entre a severidade ou a tipologia do crime e a simulação de sintomas, apesar de os reclusos que cumpriam pena há mais tempo, apresentarem maior probabilidade de ser identificados como simuladores pela

SIRS.

A avaliação da simulação em contexto prisional responde a uma série de princípios e particularidades que devem ser tidas em análise. O SIMS apresenta-se como um importante instrumento que deve ser considerado na avaliação neste contexto, considerando tanto uma amostra jovem, como adulta (Rogers, Hinds, & Sewell, 1996). Para além disto, a aplicação deste inventário revela-se útil na medida em que pode ser administrado em pequenos grupos, de forma relativamente rápida, não exigindo recursos profissionais que são, muitas vezes, escassos neste meio. Contudo, a sua utilidade tem vindo a ser contestada, sendo que em contexto prisional têm sido discutidas algumas questões. Edens e colaboradores (2007) conduziram uma investigação com uma amostra de reclusos com o objetivo específico de averiguar se este instrumento se constituía uma medida suficiente no reconhecimento e classificação da simulação. Os resultados concluíram que o teste foi capaz de detetar 69% dos sujeitos simuladores, contudo, os autores recomendam que, tratando-se de uma amostra composta por indivíduos reclusos, o SIMS não deverá ser a única medida utilizada na avaliação deste fenómeno, sob pena de aumentar a percentagem de falsos positivos. A SIRS, apesar de se tratar de uma entrevista longa, que requer algum tempo de administração, tem também vantagens na sua aplicação em meio carcerário, relacionando-se a maior com o seu formato de resposta aberta que não implica que os indivíduos escrevam ou leiam. Esta particularidade é importante dados os níveis de analfabetismo neste meio serem relativamente elevados (Norris & May, 1998). Os inventários de avaliação da personalidade, como o MMPI-2 e o PAI, têm também a sua utilização sustentada por vários estudos em contexto prisional. Contudo, a informação acerca da validade de sintomas que é possível retirar destas medidas é um pouco limitada, considerando que esse não é o seu propósito avaliativo (Vitacco & Rogers, 2005).

Mesmo os profissionais mais conservadores e reticentes quanto ao uso do conceito de simulação assumem preocupações quanto às elevadas prevalências encontradas nos vários contextos forenses (Greve et al., 2008), dado que poderão ser entre 10 a 30 vezes superiores àquelas encontradas em ambulatório de psiquiatria (Gil et al., 2017). As evidências vão no sentido de que as prevalências de relatos pouco credíveis no contexto forense variem entre 30 a 54% (Larrabee, 2003; Martin et al., 2015). Os profissionais estimam que, em amostras forenses, as prevalências se situem entre os 15 e os 18% (Rogers et al., 1994; Rogers et al., 1998). Contudo, em amostras psiquiátricas forenses, a investigação sugere prevalências que podem ultrapassar os 25% (Bocaccini et al., 2006; Lewis et al., 2002). Ainda neste sentido, a investigação tem demonstrado que é provável que os comportamentos de simulação assumam prevalências maiores em contextos penais, comparativamente a contextos civis (Mittenberg et al., 2002; Denney, 2008). Especificamente no contexto prisional, não são completamente claros os valores para a prevalência da simulação (Wall et al., 2015). Já na década de 80 foram encontradas prevalências significativas deste problema em reclusos. Num estudo de Walter e colaboradores (1988), 46% dos sujeitos que procuravam acompanhamento psicológico foram classificados como

simuladores. Pollock e colaboradores (1997) estudaram uma amostra forense composta por agressores a cumprir pena numa prisão de segurança média, e concluíram que, dos 67 reclusos, 32% terão procurado simular sintomas nas respostas aos instrumentos de avaliação (SIRS e MMPI-2). Num estudo de Norris e May (1998) realizado com uma amostra de reclusos de ambos os sexos que solicitavam acompanhamento psicológico, os resultados sugerem que entre 45 e 56% dos inquiridos foram definidos como simuladores. Mais recentemente, McDermott, Dualan e Scott (2013) encontraram prevalências de 64.5% de sujeitos simuladores numa amostra de reclusos que procuravam apoio psiquiátrico. Para além disto, ao analisarem uma amostra de indivíduos considerados inaptos para julgamento, aproximadamente 17% dos sujeitos foram classificados, na entrevista de avaliação inicial, como simuladores.

No que toca à população mais jovem, o conhecimento das prevalências e manifestações da simulação neste contexto, é relativamente mais escasso. As preocupações mais evidentes neste sentido relacionam-se, sobretudo, com a carência de medidas suficientemente válidas que permitam a avaliação da produção ou exagero de sintomas em jovens que cumprem medidas privativas de liberdade. Contudo, tanto a SIRS, como o SIMS se assumem como instrumentos úteis para a avaliação da validade de sintomas na aplicação nestes contextos. No entanto, as potencialidades do SIMS apenas sustentam a sua utilização enquanto medida de rastreio (Rogers, Hind & Sewell, 1996).

Uma questão pertinente prende-se com o modo de atuação dos profissionais de saúde mental perante resultados sugestivos de um perfil de simulação. Tratando-se esta de uma população altamente sensível à crítica e à confrontação, a devolução de informação poderá não ser uma resposta viável, para além de eventualmente poder comprometer a relação terapêutica. No entanto, poderá dar-se oportunidade ao indivíduo de responder novamente ao protocolo, argumentando que os resultados foram invalidados por alguma razão.

O ambiente prisional poderá ser propício à aprendizagem de comportamentos de simulação, uma vez que, à medida que este se torna mais restritivo e punitivo, alguns sujeitos adotam estratégias de coping no sentido da adaptação ao meio que poderão ser difíceis de distinguir de alguns comportamentos de simulação (Resnick & Knoll, 2008). A avaliação da simulação de sintomas em meio prisional constitui um enorme desafio para os profissionais de saúde mental que aqui atuam, dada a responsabilidade e a gravidade das consequências que poderão advir de erros a este nível. A incorreta classificação de um sujeito como simulador pode causar a negação não intencional de tratamentos de saúde, denegrir a sua imagem e prejudicar o relacionamento do clínico com o paciente (Conroy & Kwartner, 2006; McDermott & Sokolov, 2009; McDermott et al., 2013). Por outro lado, a não identificação de um simulador, para além de comprometer todo o protocolo avaliativo, contribuirá para que seja posta em causa a credibilidade da avaliação e do perito perante o sistema de justiça (Barber-Rioja et al., 2016).

Concluindo, os profissionais de saúde mental que procuram avaliar a simulação na população reclusa, deverão considerar as limitações da definição

proposta pelo DSM, utilizando sempre estratégias e instrumentos específicos e suficientemente validados para a avaliação deste fenómeno.

V – Questões éticas na avaliação da Simulação de Sintomas

Existe uma quantidade considerável de concepções erróneas associadas ao conceito e à prevalência da simulação, tanto no que concerne ao contexto clínico, como forense. Estes estereótipos traduzem-se numa variedade de questões éticas que devem ser exploradas e tidas em consideração.

Há várias razões pelas quais um indivíduo pode produzir ou exagerar sintomas, nomeadamente: o contexto específico em que decorre a avaliação, chamar a atenção do clínico para determinada queixa, a presença de determinados traços de personalidade (Iverson, 2007), e, claro, a evidente intenção de produzir ou exagerar sintomas. Os instrumentos de avaliação da simulação, embora fundamentais em qualquer processo de avaliação, não determinam as razões pelas quais o sujeito está a reportar excessivamente sintomatologia. Como tal, havendo evidência desse relato exagerado, os examinadores deverão enquadrar o perfil de resultados nos dados provenientes da entrevista clínica, da observação direta do comportamento, e de fontes de informação colaterais, uma vez que as causas para o relato exagerado de sintomas poderão ter origem psiquiátrica, neurológica ou desenvolvimental (Drogin & Williams, 2018). A avaliação detalhada da simulação de sintomas (e dos outros estilos de resposta) é, de facto, crucial. Se um sujeito for bem-sucedido na sua intenção de simular e por consequência, obtiver o ganho secundário desejado, as consequências poderão ser graves. No contexto forense, por exemplo, um indivíduo que tenha cometido um crime severo, mas tenha falsificado sintomatologia, poderá não ser devidamente punido, ou ser reinserido mais cedo na sociedade, sem que sobre ele tenha sido realizada uma intervenção ajustada (Carr, 2015). Assim, um diagnóstico de simulação, deverá ser sempre baseado no julgamento clínico do examinador a par dos resultados das medidas cientificamente validadas (Conroy & Kwarter, 2006; Merckelbach, Dandachi-FitzGerald, Helvoort & Otgaar, 2019; Seron, 2014).

Alguns profissionais, por falta de conhecimento ou experiência, assumem ainda que este se trata de um fenómeno raro, optando assim por não incluir a sua avaliação específica nos protocolos elaborados. Contudo, as prevalências não corroboram tal premissa em nenhum contexto de avaliação, devendo sempre ser considerado que, quando o resultado de uma avaliação tem consequências importantes para o sujeito, o domínio específico da simulação de sintomas deverá sempre ser apreciado, sob pena de tal omissão transparecer alguma negligência por parte do profissional (Drogin & Williams, 2018).

O relato sintomático excessivo é frequentemente traduzido num perfil de simulação, cuja conotação associada é tendencialmente negativa e associada a desonestidade e traços antissociais (Iverson & Binder, 2000), não devendo esta ligação ser diretamente assumida. Apesar de a simulação ou o relato exagerado de sintomas assumirem prevalências mais significativas em contexto forense, tal não significa que todas as pessoas envolvidas em pedidos

de indemnização, reformas antecipadas, ou que buscam ressarcimento pelas suas incapacidades, evidenciem comportamentos de esforço suspeito (Simões & Sousa, 2008). De facto, Pollock e colaboradores (1997) procuraram estudar os traços de personalidade mais prevalentes em sujeitos simuladores e os resultados sugerem que a simulação de psicopatologia não está necessariamente ligada à presença de traços antissociais de personalidade. Pelo contrário, o grupo de simuladores apresentou maior prevalência de traços de dependência e ansiedade.

Alguns indivíduos com patologias crónicas poderão ser mais difíceis de avaliar, sendo necessária alguma cautela nas conclusões que daí se inferem. Alguns pacientes poderão reconhecer a doença como parte da sua identidade, ou sentirem-se de tal forma assoberbados pela sintomatologia que tendem a comunicar queixas/sintomas mais severos do que aqueles que realmente sentem (Rogers & Payne, 2006; citados em Drogin & Williams, 2018). Esta comunicação excessiva de sintomas em nada se relaciona com um perfil de simulação, uma vez que o propósito destes relatos não se prende com motivações externas, mas possivelmente com a necessidade de chamar a atenção para determinada queixa (Gil, Caldeira & Barreto, 2017), ou simplesmente porque a patologia se torna parte do sujeito e este sente necessidade de a evidenciar.

A opção por enveredar por um perfil de simulação num determinado momento não implica que esse seja sempre o perfil assumido pelo sujeito ao longo do percurso, i.e., o facto de ter produzido/exagerado sintomatologia num dado momento não implica que o faça sempre, pelo que tal não deverá ser rotulador (Bender & Frederick, 2018). Apesar de, na maioria dos casos, esse relato se traduzir realmente num perfil de simulação, esta relação não deverá ser rigorosamente assumida. Há uma clara tendência para atribuir a presença de simulação de sintomas a fatores individuais, subvalorizando fatores contextuais. A simulação não está circunscrita a contextos de avaliação psicológica. Deturpar informação, seja no sentido de enaltecer características/feitos pessoais, seja para exagerar queixas ou sintomas, são comportamentos relativamente comuns. Tal poderá ocorrer, por exemplo, num cenário de absentismo escolar ou no sentido de obter baixa médica (Iverson, 2003). De facto, mesmo no que concerne especificamente ao âmbito da avaliação forense, não é aconselhável para a validade da avaliação que todas e quaisquer informações sejam minuciosamente explicadas no consentimento informado, caso contrário, seria expectável que o sujeito procurasse alterar as suas respostas. Por questões éticas, o sujeito deverá ser informado dos domínios que serão alvo de avaliação (cognição, personalidade, funcionamento emocional, ...), incluindo que existirão parâmetros que se destinam à avaliação da honestidade e consistência das respostas. Fornecer informações mais específicas neste sentido poderá ser contraproducente e retirar alguma da autenticidade às respostas do sujeito (Bush et al., 2005; Drogin & Williams, 2018). Ora, logo nos primeiros momentos do processo muitas informações são omitidas ou não exploradas, o que poderá minar a transparência da relação examinador-examinando (Bender & Frederick, 2018).

Ainda neste sentido, importa refletir acerca da comunicação de um eventual perfil inválido, uma vez que os profissionais se deparam frequentemente com o dilema de comunicar ou não ao indivíduo que se sabe que o seu relato não foi autêntico. Um estudo realizado com 209 (neuro)psicólogos concluiu que 98% dos profissionais concordava que o *feedback* da avaliação deveria incluir alguma informação acerca da eventual invalidação do perfil, sendo que 75% afirma que deverão ser fornecidas explicações detalhadas e o impacto que a invalidação terá na interpretação do perfil. Contudo, não existe consenso entre os profissionais na forma como esse *feedback* deverá ser fornecido e que informações específicas deverão constar. Não existindo, até à data, um modelo que providencie *guidelines* específicas, a forma como esta informação é devolvida ao sujeito varia entre profissionais (Martin & Schroeder, 2020).

O facto de não ser possível medir com rigor as motivações por detrás de um perfil de simulação (Rogers et al., 1993) fomenta a existência de muitos estereótipos, i.e., crenças baseadas na ideia destes sujeitos enquanto indivíduos saudáveis, manipuladores e mentirosos (Grossman & Wasyliv, 1988). Esta ambiguidade na caracterização do conceito e o consequente surgimento de rótulos depreciativos tem vindo a apoiar várias propostas no sentido da supressão do termo “simulação” em relatórios, cuja evidência do fenómeno não é rigorosa o suficiente para sustentar o seu diagnóstico (Berry & Nelson, 2010). A conceptualização alternativa do termo visa que este seja efetivamente encarado como um comportamento comum e, muitas vezes, adaptativo (Heilbronner et al., 2009; Rogers, 1990), e não como um ato imoral (Simões, 2012). Assim, a inclusão da palavra “simulação” nos relatórios deve ser feita com muita cautela e sustentada por instrumentos rigorosos e válidos, sob pena de ter consequências gravosas e negativas para o sujeito. Caso o examinador não esteja perfeitamente seguro de que se trata de um caso de simulação de sintomas, deverá optar por utilizar conceitos como “falsificar”, “exagerar” ou “dissimular” (Rogers, 2018c). Berry e Nelson (2010) acrescentam que à descrição destes perfis poderá ser adicionado um grau no sentido de caracterizar a produção/exagero de sintomatologia – fraco, moderado, severo. Na eventualidade de existir uma forte inclinação para a presença de simulação, mas não exista um grau de certeza suficiente, o perito deverá nomear as inconsistências na história relatada pelo sujeito, ou as eventuais incongruências nos resultados psicométricos, mencionando que não se encontra afastada a hipótese da simulação, embora admitindo não ter condições suficientes para uma conclusão definitiva sobre tal possibilidade (Gil et al., 2017). Neste sentido, a investigação tem demonstrado a preocupação dos (neuro)psicólogos face à utilização do termo “simulação” nos relatórios. De acordo com um estudo realizado por Martin e colaboradores (2015) com uma amostra de 316 profissionais, aproximadamente 90% afirma incluir raramente ou nunca este termo nos relatórios, o que sugere que os profissionais se encontram cada vez mais reticentes quanto a concluir acerca das intenções ou motivações subjacentes a um relato falso ou exagerado de sintomatologia. Especificamente, apenas 24% dos examinadores forenses refere utilizar a palavra “simulação” em casos em que há suspeita deste

fenómeno.

Concluindo, importa realçar a ideia de que apesar de todos os sujeitos simuladores relatarem excessivamente sintomas, nem todos os indivíduos que o fazem devem ser classificados como simuladores (Iverson, 2007; Merckelbach et al., 2019). Um diagnóstico erróneo de simulação poderá ter consequências graves para o indivíduo, nomeadamente porque afetará a sua credibilidade e os resultados nos testes podendo, em casos mais extremos, limitar-lhe o acesso a direito e tratamentos clínicos (Drogin & Williams, 2018), ou impedir que sejam aplicadas medidas legais justas (Kucharski, Toomey, Fila, & Duncan, 2007).

VI - *Self-Report Symptom Inventory* (SRSI) em contexto prisional: Proposta de estudo com *design* de grupos conhecidos

O desenvolvimento do SRSI dá-se no sentido de responder às limitações apontadas aos vários TVS, ao mesmo que tempo que visa complementar a avaliação da simulação (Pinheiro, 2019; van Helvoort et al., 2019). Esta escassez de TVS adaptados e validados é particularmente notória na população portuguesa, sobretudo considerando as mais recentes críticas apontadas ao SIMS, medida mais utilizada para este propósito.

Uma das limitações apontadas às várias investigações recentes desenvolvidas com o SRSI prende-se com o facto de ainda não terem sido realizados estudos com o *design* de grupos conhecidos, que incluam grupos compostos por sujeitos previamente classificados como simuladores (Merten et al., 2016). Para além disto, considerando a escassez de instrumentos suficientemente válidos no âmbito da avaliação da validade de sintomas para a população prisional portuguesa, seria oportuno o estudo e a adaptação do SRSI neste meio específico.

No sentido de complementar os estudos que, até ao momento, visam a adaptação deste instrumento ao contexto português (Domingues, 2019; Dwarkadas, 2018; Pinheiro, 2019), será apresentada, neste ponto, uma proposta de estudo que visa a validação do SRSI especificamente para o contexto prisional português, através da implementação do *design* de grupos conhecidos, cujas potencialidades têm sido reconhecidas. Assim, e considerando as recomendações à utilização deste *design*, a presente investigação contemplaria a constituição de dois grupos: um grupo de sujeitos simuladores, e outro de sujeitos não simuladores, i.e., respondentes honestos. Após esta classificação, todos os sujeitos responderiam ao SRSI no sentido de comparar os resultados e classificação fornecida por esta medida, com aqueles previamente considerados para a constituição dos grupos. Para que a utilização deste instrumento fosse validada, os sujeitos anteriormente classificados como simuladores, teriam também de o ser de acordo com o SRSI.

Não se tratando este *design* de uma medida extensivamente usada, é imperativo atender a uma série de recomendações no sentido de garantir a fiabilidade das conclusões da investigação. Considerando as limitações das medidas de autorrelato usadas no âmbito da avaliação da validade de sintomas, importa que, de forma a garantir a validade de um perfil de

simulação, sejam utilizados vários instrumentos neste sentido. Assim, importa atender às recomendações de De Marchi e Balboni (2018) para a constituição dos grupos que seguem no sentido de serem consideradas mais do que uma medida na classificação dos sujeitos e posterior inclusão nos grupos honesto/simulador. Assim, seriam utilizadas duas medidas de autorrelato na constituição dos referidos grupos: o SIMS e a EVS-2. Apesar das fragilidades apontadas ao SIMS, a sua inclusão dá-se pelo facto de este ser um instrumento extensivamente usado no âmbito da avaliação da simulação em meio prisional, contando com diversos estudos em Portugal com resultados conhecidos (Simões, Almiro, Mota et al., 2017). A EVS-2, para além de se tratar de uma medida desenvolvida em Portugal, é um instrumento útil a ser utilizado de forma complementar ao SIMS, dado que são instrumentos baseados nos mesmos constructos de simulação ou exagero de sintomas (Simões, Almiro, Mota et al., 2017). Para além destas medidas, seria oportuna a administração de um inventário destinado ao controlo e avaliação da psicopatologia, como o Inventário de Sintomas Psicopatológicos (BSI; Derogatis, 1982; Canavarro, 2007), para que reclusos com real psicopatologia não fossem considerados simuladores e incluídos na análise. Também seria útil a administração de um breve questionário de forma a compreender as características sociodemográficas, jurídico-penais e clínicas da amostra.

No que concerne aos objetivos específicos do estudo, importa analisar vários pontos: em primeiro lugar, comparar-se-iam os resultados do SRSI com os resultados de um grupo de simuladores, previamente classificados como tal pelo SIMS e pela EVS-2; de seguida, seria conveniente uma análise da influência de variáveis jurídico-penais (situação penal, antecedentes criminais, duração da pena) e clínicas (historial de toxicodependência, toma de medicação, internamentos prévios, acompanhamento psicológico/psiquiátrico, historial de comportamentos autolesivos) nos resultados do SRSI; por fim, comparar-se-iam os resultados obtidos neste mesmo estudo, com aqueles que Pinheiro (2019) obteve no seu estudo com o SRSI também em contexto prisional português, e com os resultados de um estudo alemão, igualmente neste âmbito (Huhnt, 2013).

No entanto, considerando os dados de investigação acerca da simulação e exagero de sintomas no meio prisional, algumas hipóteses poderiam ser estabelecidas e, posteriormente, testadas. Por um lado, seria expectável que o SRSI, enquanto instrumento destinado à avaliação da simulação de sintomas, identificasse como simuladores os sujeitos que previamente foram reconhecidos como tal pelo SIMS e pela EVS-2. Neste sentido, importa considerar que, de acordo com Rogers (2018a), os casos intermédios (sujeitos identificados como simuladores em apenas um dos instrumentos) ou prováveis (resultados próximos dos pontos de corte) de simulação não deveriam ser incluídos na constituição dos grupos, dado que se pretende que a classificação de simulação seja rigorosa, diminuindo a probabilidade de falsos positivos. Assim, para que um sujeito fosse classificado como simulador deveria ser identificado como tal tanto pela EVS-2, como pelo SIMS, utilizando-se, neste último, um ponto de corte >24 , uma vez que se pretende que o instrumento identifique com exatidão os sujeitos

simuladores e não simuladores, não persistindo questões dúbias quanto à classificação. Uma outra hipótese a ser estabelecida relaciona-se com o facto de, no domínio da validade convergente, poderem ser esperadas correlações positivas entre os resultados do SIMS e da EVS-2, e a escala de Pseudosintomas do SRSI, uma vez que se tratam de medidas destinadas à avaliação de construtos semelhantes. Neste sentido, seria também expectável que os resultados espelhassem relações positivas entre os resultados globais do BSI e a escala de Sintomas Genuínos do SRSI. Por fim, e tal como sugere Denney (2007), seria provável que a amostra de reclusos preventivos obtivesse valores globais mais elevados nos instrumentos destinados à avaliação da simulação, comparativamente à amostra composta por reclusos condenados, dado que, não havendo ainda conclusão do processo penal, estes indivíduos procurariam a atenuação da culpa e, conseqüentemente, da pena de prisão. No sentido de testar esta última hipótese e visando obter uma amostra representativa da população reclusa, sugere-se que fossem selecionados indivíduos a cumprir pena, por exemplo, no Estabelecimento Prisional de Coimbra, uma unidade central cuja população prisional se encontra, maioritariamente, em situação de condenação, e no Estabelecimento Prisional de Aveiro, onde se encontram reclusos preventivos, na sua maioria.

No sentido de salvaguardar a integridade e credibilidade da investigação, a amostra seria selecionada com base em alguns critérios. Em primeiro lugar, seriam excluídos reclusos analfabetos, dado que são requeridas competências linguísticas e de leitura básicas na resposta aos instrumentos psicométricos em causa. Ainda neste âmbito, no sentido de neutralizar eventuais barreiras linguísticas que distorcessem a interpretação dos itens, não seriam incluídos na amostra indivíduos com evidentes dificuldades na língua portuguesa. Por fim, seriam excluídos deste estudo reclusos com evidentes problemas cognitivos ou psicopatologia associada, sendo este último ponto avaliado através dos resultados obtidos no BSI.

Como analisado anteriormente, a população reclusa apresenta diversas razões para simular ou exagerar sintomatologia. Como tal, é necessário que os instrumentos que se destinam à avaliação neste âmbito, se encontrem suficientemente validados e aferidos para este contexto específico. O SRSI apresenta-se como uma medida com várias potencialidades neste sentido, o que sustenta a importância da continuidade das investigações orientadas para a sua validação no meio prisional.

VII – Conclusão

Habitualmente, assume-se o esforço dos indivíduos no sentido de manifestarem o melhor desempenho nas respostas aos instrumentos de avaliação (e.g., testes de inteligência, testes neurocognitivos), e a honestidade nas respostas a técnicas de autorrelato (e.g., entrevistas, inventários de sintomas psicopatológicos, questionários de personalidade). Esta amigável relação direta entre funcionamento psicológico real e desempenhos e respostas verbais tem sido contraditada, uma vez que se a investigação e a prática forenses evidenciam que o perfil de resultados pode ser invalidado devido ao pouco esforço ou intenção clara do sujeito em manipular os

resultados. A investigação e a prática forenses têm demonstrado a necessidade da avaliação da validade de sintomas (e desempenhos) e a indispensabilidade da sua integração nos protocolos de avaliação forense (Bush et al., 2005; Heilbronner et al., 2009; Martin et al., 2015).

Particularmente no que toca à simulação ou exagero de sintomas, as elevadas taxas de prevalência evidenciam a maior probabilidade deste problema ser encontrado em contextos forenses, e entre estes o contexto médico-legal, devido ao facto de que muitos indivíduos podem ser tentados a modificar o relato de sintomas a fim de obter incentivo externo, como uma compensação monetária ou diminuição da responsabilidade criminal. No entanto, é redutor atribuir a maioria da invalidação dos protocolos à presença de simulação, uma vez que este estilo de resposta não esgota os cenários nem as explicações para o problema da falta de validade dos resultados da avaliação (neuro)psicológica.

Considerando a necessidade de inclusão de vários instrumentos específicos de identificação da simulação e exagero de sintomas nos protocolos de avaliação, revela-se fundamental a sua validação nos diferentes contextos e com base em diferentes métodos. Dos vários tipos de *designs* utilizados para a validação de instrumentos de avaliação, o *design* de grupos conhecidos assume evidentes potencialidades no âmbito da validade de sintomas, particularmente no que concerne à simulação e exagero de sintomas. A utilidade do *design* de grupos conhecidos prende-se, sobretudo, com as vantagens ao nível da validade externa.

A revisão de estudos apresentada neste trabalho pretende contribuir para a melhor compreensão do *design* de grupos conhecidos e a sua aplicação na avaliação da simulação de sintomas em vários contextos forenses, no sentido de identificar instrumentos, estratégias e prevalências deste problema. A conclusão de relatos deturpados (falta de validade ou validade questionável de desempenhos e respostas) levanta questões, não apenas técnicas e científicas, mas também, e de modo indissociável, dilemas éticos que requerem uma análise simultaneamente rigorosa e imparcial das consequências.

Bibliografia

Almeida, D. (2015). *Comportamento de simulação e exagero de sintomas em contexto médico-legal: Novos Estudos de validação do Structured Inventory of Malingered Symptomatology (SIMS)* (Dissertação de Mestrado). Coimbra: Faculdade de Psicologia e Ciências da Educação da Universidade de Coimbra.

American Psychiatric Association (2014). *DSM-5: Manual de Diagnóstico e Estatística de Perturbações Mentais*. Lisboa: Climepsi Editores.

Archer, R.P., Buffington-Vollum, J.K., Stredny, R.V., & Handel, R.W. (2006). A Survey of Psychological Test Use Patterns Among Forensic Psychologists. *Journal of Personality Assessment*, 87(1), 84–94. doi:10.1207/s15327752jpa8701_07

Barber-Rioja, V., Zottoli, T. M., Kucharski, L. T., & Duncan, S. (2009). The Utility of the MMPI-2 Criminal Offender Infrequency (Fc) Scale in the Detection of Malingering in Criminal Defendants. *International Journal of Forensic Mental Health*, 8(1), 16–24. doi:10.1080/14999010903014689

Bass, C., & Halligan, P. (2014). Factitious disorders and malingering: Challenges for clinical assessment and management. *The Lancet*, 383(9926), 1422–1432. doi:10.1016/s0140-6736(13)62186-8

Bender, S., & Frederick, R. (2018). Neuropsychological Models of Feigned Cognitive Deficits. In Richard Rogers & Scott D. Bender (Eds.). *Clinical Assessment of Malingering and Deception* (4th Ed., pp. 42-60). New York: Guilford.

Berry, D. T. R., & Nelson, N. W. (2010). DSM-5 and Malingering: A Modest Proposal. *Psychological Injury and Law*, 3(4), 295–303. doi:10.1007/s12207-010-9087-7

Bianchini, K. J., Etherton, J. L., Greve, K. W., Heinly, M. T., & Meyers, J. E. (2008). *Classification Accuracy of MMPI-2 Validity Scales in the Detection of Pain-Related Malingering*. *Assessment*, 15(4), 435–449. doi:10.1177/1073191108317341

Bigler, E. D. (2012). Symptom Validity Testing, Effort, and Neuropsychological Assessment. *Journal of the International Neuropsychological Society*, 18(04), 632–

642. doi:10.1017/s1355617712000252

Boccaccini, M. T., & Hart, J. R. (2018). Response Style on the Personality Assessment Inventory and Other Multiscale Inventories. In Richard Rogers & Scott D. Bender (Ed.). *Clinical Assessment of Malingering and Deception* (4th ed., pp. 280-300). New York: Guilford.

Boccaccini, M. T., Murrie, D. C., & Duncan, S. A. (2006). Screening for malingering in a criminal-forensic sample with the Personality Assessment Inventory. *Psychological Assessment*, 18(4), 415-423. doi: 10.1037/1040-3590.18.4.415

Boskovic, I., Merckelbach, H., Merten, T., Hope, L., & Jelicic, M. (2019). The Self Report Symptom Inventory as an Instrument for Detecting Symptom Over-Reporting: An Exploratory Study With Instructed Simulators. *European Journal of Psychological Assessment*. Advance online publication. doi:10.1027/1015-5759/a000547

Bush, S. S., Ruff, R. M., Troster, A. I., Barth, J. T., Koffler, S. P., Pliskin, N. H., & Silver, C. H. (2005). Symptom validity assessment: Practice issues and medical necessity - NAN policy & planning committee. *Archives of Clinical Neuropsychology*, 20, 419-426. doi: 10.1016/j.acn.2005.02.002

Carr, C. M. K. (2015). *The assessment of malingering within forensic populations*. Loma Linda University Electronic Theses, Dissertations & Projects. Retrieved from <https://scholarsrepository.llu.edu/etd/>

Cornell, D. G., & Hawk, G. L. (1989). Clinical presentation of malingers diagnosed by experienced forensic psychologists. *Law and Human Behavior*, 13(4), 375– 383. doi:10.1007/bf01056409

Cernovsky, Z. Z., Ferrari, J. J. R., & Mendonça, J. D. (2019). Pseudodiagnoses of Malingering of Neuropsychological Symptoms in Survivors of Car Accidents by the Structured Inventory of Malingered Symptomatology. *Archives of Psychiatry and Behavioral Sciences*, 2(1), 55-65.

Cernovsky, Z., Mendonça, J., Oyewumi, L., Ferrari, J., Sidhu, G., Campbell, R. (2019). Content Validity of the Psychosis Subscale of the Structured Inventory of Malingered Symptomatology (SIMS). *International Journal of Psychology and Cognitive Science*, 5(3), 121-127.

Clegg, C., Fremouw, W., & Mogge, N. (2009). Utility of the Structured Inventory of Malingered Symptomatology (SIMS) and the Assessment of

Depression Inventory (ADI) in screening for malingering among outpatients seeking to claim disability. *Journal of Forensic Psychiatry & Psychology*, 20(2), 239–254. doi:10.1080/14789940802267760

Conroy, M. A. & Kwartner, P. P. (2006). Malingering. *Applied Psychology in Criminal Justice*, 2(3), 29-51.

Dandachi-FitzGerald, B., Duits, A. A., Leentjens, A. F. G., Verhey, F. R. J., & Ponds, R. W. H. M. (2019). Performance and Symptom Validity Assessment in Patients with Apathy and Cognitive Impairment. *Journal of the International Neuropsychological Society*, 1–8. doi:10.1017/s1355617719001139

De Marchi B. & Balboni G. (2018). Detecting malingering mental illness in forensics: Known-Group Comparison and Simulation Design with MMPI-2, SIMS and NIM. *PeerJ*. doi: 10.7717/peerj.5259

Denney, R. L. (2007). Assessment of malingering in criminal forensic neuropsychological settings. In K. B. Boone (Ed.), *Assessment of feigned cognitive impairment: A neuropsychological perspective* (pp. 428-452). New York: The Guilford.

Denney, R. L. (2008). Negative response bias and malingering during neuropsychological assessment in criminal forensic setting. In R. L. Denney & J. P. Sullivan (Eds.), *Criminal neuropsychology in the criminal forensic setting* (pp. 91-133). New York: Guilford.

Domingues, M. (2019). *Self-Report Symptom Inventory (SRSI): Validação através de um estudo de simulação análogo* (Dissertação de Mestrado). Coimbra: Faculdade de Psicologia e Ciências da Educação da Universidade de Coimbra.

Drogin, E. Y., & Williams, C. S. (2018). Malingering: Considerations in Reporting and Testifying about Assessment Results. In Richard Rogers & Scott D. Bender. *Clinical Assessment of Malingering and Deception* (4th Ed, pp. 514-529). New York: Guilford.

Dwarkadas, A. (2018). *Self-Report Symptom Inventory (SRSI): Adaptação, Validação e Normas com base em Estudos na Comunidade* (Dissertação de Mestrado). Coimbra: Faculdade de Psicologia e de Ciências da Educação da Universidade de Coimbra.

Edens, J. F., Buffington, J. K., & Tomicic, T. L. (2000). An investigation of the relationship between psychopathic traits and malingering

on the Psychopathic Personality Inventory. *Assessment*, 7, 281-296. doi: 10.1177/107319110000700307

Edens, J. F., Poythress, N. G., & Watkins-Clay, M.M. (2007). Detection of malingering in psychiatric unit and general population prison inmates: A comparison of the PAI, SIMS, and SIRS. *Journal of Personality Assessment*, 88, 33-42. doi: 10.1080/00223890709336832

Faust, D. (1995). *The Detection of Deception. Neurologic Clinics*, 13(2), 255–265. doi:10.1016/s0733-8619(18)30044-6

Ferreira, S. (2017). *Escala de Validade de Sintomas – Versão 2 (EVS-2): Estudos de validade concorrente na comunidade e em contexto medico-legal* (Dissertação de Mestrado). Coimbra: Faculdade de Psicologia e de Ciências da Educação da Universidade de Coimbra.

Geurten, M., Meulemans, T., & Seron, X. (2018). Detecting over-reporting of symptoms: The French version of the self-report symptom inventory. *The Clinical Neuropsychologist*, 1–18. doi:10.1080/13854046.2018.1524027

Giger, P., & Merten, T. (2019). Equivalence of the German and the French versions of the Self-Report Symptom Inventory. *Swiss Journal of Psychology*, 78, 5–13. doi:10.1024/1421-0185/a000218

Gil, N. P., Caldeira, S., & Barreto, H. (2017). Simulação, dissimulação e perturbações factícias. In Fernando Vieira, Ana C. Cabral & Carlos B. Saraiva (Coords.), *Manual de Psiquiatria Forense* (pp. 277-305). Lisboa: Pactor.

Green, P., & Merten, T. (2013). Noncredible explanations of noncredible performance on symptom validity tests. In D. A. Carone & S. S. Bush. (Eds.), *Mild traumatic brain injury: Symptom validity assessment and malingering* (pp. 73-100). New York: Springer.

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. doi:10.1080/13803390500263550

Greve, K. W., Ord, J., Curtis, K. L., Bianchini, K. J., & Brennan, A. (2008). *Detecting Malingering in Traumatic Brain Injury and Chronic Pain: A Comparison of Three Forced-Choice Symptom Validity Tests*. *The Clinical Neuropsychologist*, 22(5), 896–918. doi:10.1080/13854040701565208

Greve, K. W., Ord, J. S., Bianchini, K. J., & Curtis, K. L. (2009). Prevalence of Malingering in Patients With Chronic Pain Referred for Psychologic Evaluation in a Medico-Legal Context. *Archives of Physical Medicine and Rehabilitation*, *90*(7), 1117–1126. doi: 10.1016/j.apmr.2009.01.018

Grossman, L. S., & Wasyliv, O. E. (1988). A Psychometric Study of Stereotypes: Assessment of Malingering in a Criminal Forensic Group. *Journal of Personality Assessment*, *52*(3), 549–563. doi: 10.1207/s15327752jpa5203_18

Hanley, J. R., Baker, G. A., & Ledson, S. (1999). Detecting the Faking of Amnesia: A Comparison of the Effectiveness of Three Different Techniques for Distinguishing Simulators from Patients with Amnesia. *Journal of Clinical and Experimental Neuropsychology*, *21*(1), 59–69. doi: 10.1076/jcen.21.1.59.936

Hattie, J. & Cooksey, R. (1984). Procedures for Assessing the Validities of Tests Using the “Known-Groups” Method. *Applied Psychological Measurement*, *8*(3), 295–305. doi:10.1177/014662168400800306

Heilbronner, R. L., Sweet, J. J., Morgan, J. E., Larrabee, G. J., Millis, S. R., & Conference Participants. (2009). American Academy of Clinical Neuropsychology Consensus Conference Statement on the Neuropsychological Assessment of Effort, Response Bias, and Malingering. *The Clinical Neuropsychologist*, *23*(7), 1093–1129. doi:10.1080/13854040903155063

Heilbrun, K. (1992). The role of psychological testing in forensic assessment. *Law and Human Behavior*, *16*, (257–272). doi: 10.1007/BF01044769

Heubrock, D., & Petermann, F. (1998). Neuropsychological assessment of suspected malingering: Research results, evaluation techniques, and further directions of research and application. *European Journal of Psychological Assessment*, *14*, 211–225. doi: 10.1027/1015-5759.14.3.211

Huhnt, B. (2013). Stone cold manipulation or just an excuse? Psychopathy and blame attribution as possible explanations for faking symptoms in a forensic setting. (Unpublished master’s thesis). Maastricht University, Faculty of Psychology and Neuroscience, Maastricht, The

Netherlands.

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., & Binder, L. M. (2000). Detecting Exaggeration and Malingering in Neuropsychological Assessment. *Journal of Head Trauma Rehabilitation, 15*(2), 829–858. doi:10.1097/00001199-200004000-00006

Jelicic, M., Hessels, A., & Merckelbach, H. (2006). Detection of feigned psychosis with the Structured Inventory of Malingered Symptomatology (SIMS): A study of coached and uncoached simulators. *Journal of Psychopathology and Behavioral Assessment, 28*, 19-22. doi: 10.1007/s10862-006-4535-0.

Kucharski, L. T., Toomey, J. P., Fila, K., & Duncan, S. (2007). Detection of Malingering of Psychiatric Disorder With the Personality Assessment Inventory: An Investigation of Criminal Defendants. *Journal of Personality Assessment, 88*(1), 25-32. doi:10.1080/00223890709336831

Larrabee, G. (2003). Detection of Malingering Using Atypical Performance Patterns on Standard Neuropsychological Tests. *The Clinical Neuropsychologist, 17*(3), 410– 425. doi:10.1076/clin.17.3.410.18089

Larrabee, G. (2007). Introduction: Malingering, research designs, and base rates. In G. J. Larrabee (Ed.), *Assessment of malingered neuropsychological deficits* (pp. 3–13). New York: Oxford University Press.

Larrabee, G. (2008). Aggregation across Multiple Indicators Improves the Detection of Malingering: Relationship to Likelihood Ratios. *The Clinical Neuropsychologist, 22*(4), 666-679. doi: 10.1080/13854040701494987

Larrabee, G. (2014). Test Validity and Performance Validity: Considerations in Providing a Framework for Development of an Ability-Focused Neuropsychological Test Battery. *Archives of Clinical Neuropsychology, 29*(7), 695-714. doi: 10.1093/arclin/acu049

Lewis, J. L., Simcox, A. M., & Berry, D. T. R. (2002). Screening for feigned psychiatric symptoms in a forensic sample by using the MMPI-2 and the Structured Inventory of Malingered Symptomatology. *Psychological Assessment, 14*(2), 170-176. doi: 10.1037/1040-3590.14.2.170

Lilienfeld, S. O., Thames, A. D., & Watts, A. L. (2013). Symptom - Validity Testing: Unresolved Questions, Future Directions. *Journal of*

Experimental Psychopathology, 4(1), 78-87. doi:10.5127/jep.028312

Lynch, W. (2004). Determination of Effort Level, Exaggeration, and Malingering in Neurocognitive Assessment. *Journal of Head Trauma Rehabilitation*, 19(3), 277-283. doi: 10.1097/00001199-200405000-00008

Martin, P. K., Schroeder, R. W., Olsen, D. H., Maloy, H., Boettcher, A., Ernst, N., & Okut, H. (2019). A systematic review and meta-analysis of the Test of Memory Malingering in adults: Two decades of deception detection. *The Clinical Neuropsychologist*, 1-32. doi:10.1080/13854046.2019.1637027

McDermott, B. (2012). Psychological Testing and the Assessment of Malingering. *Psychiatric Clinics of North America*, 35(4), 855-876. doi: 10.1016/j.psc.2012.08.006

McDermott, B., & Sokolov, G. (2009). Malingering in a correctional setting: The use of the Structured Interview of Reported Symptoms in a jail sample. *Behavioral Sciences & the Law*, 27, 753-765. doi:10.1002/bsl.892

McDermott, B. E., Dualan, I. V., & Scott, C. L. (2013). Malingering in the correctional system: Does incentive affect prevalence? *International Journal of Law and Psychiatry*, 36(3-4), 287-292. doi: 10.1016/j.ijlp.2013.04.013

McGrath, R. E., Mitchell, M., Kim, B. H., & Hough, L. (2010). Evidence for response bias as a source of error variance in applied assessment. *Psychological Bulletin*, 136, 450-470. doi: 10.1037/a0019216

McLaughlin, J. L., & Kan, L. Y. (2014). Test usage in four common types of forensic mental health assessment. *Professional Psychology: Research and Practice*, 45(2), 128-135. <https://doi.org/10.1037/a0036318>

Merckelbach, H. & Smith, G. (2003). Diagnostic accuracy of the Structured Inventory of Malingered Symptomatology (SIMS) in detecting instructed malingering. *Archives of Clinical Neuropsychology*, 18(2), 145-152. doi:10.1016/s0887-6177(01)00191-3

Merckelbach, H., Dandachi-FitzGerald, B., Van Helvoort, D., Jelicic, M., & Otgaar, H. (2019). When Patients Overreport Symptoms: More Than Just Malingering. *Current Directions in Psychological Science*, 28(3), 321-326. doi:10.1177/0963721419837681

Merten T, & Merckelbach H. (2013) Introduction to malingering

research and symptom validity assessment. *Journal of Experimental Psychopathology*, 4(1), 3-5. doi: 10.1177/204380871300400102

Merten, T., Dandachi-Fitzgerald, B., Hall, V., Schmandd, B. A., Santamaría, P., & González-Ordi, H. (2013). Symptom validity assessment in European countries: Development and state of the art. *Clínica y Salud*, 24(3), 129–138. doi:10.1016/s1130-5274(13)70014-8

Merten, T., Merckelbach, H., Giger, P., Stevens, A. (2016). The Self-Report Symptom Inventory (SRSI): A new instrument for the assessment of distorted symptom endorsement. *Psychological Injury and Law*, 9, 102-111. doi: 10.1007/s12207-016- 9257-3

Morey, L. C., & Lowmaster, S. E. (2010). *Personality Assessment Inventory*. The Corsini Encyclopedia of Psychology. doi:10.1002/9780470479216.corpsy0663.

Nichols, D. S. (2011). *Essentials of MMPI-2 Assessment* (2nd ed.). New York: Wiley.

Norris, M. P., & May, M. C. (1998). Screening for malingering in a correctional setting. *Law and Human Behavior*, 22, 315–323. doi:10.1023/a:1025706606774

Pinheiro, D. (2019). *Self-Report Symptom Inventory (SRSI): Estudo de validação em uma amostra de reclusos* (Dissertação de Mestrado). Coimbra: Faculdade de Psicologia e Ciências da Educação da Universidade de Coimbra.

Pollock, P., Quigley, B., Worley, K., & Bashford, C. (1997). Feigned mental disorder in prisoners referred to forensic mental health services. *Journal of Psychiatric and Mental Health Nursing*, 4, 9–15. doi: 10.1111/j.1365-2850.1997.tb00171.x

Resnick, P. J., & Knoll, J. L. (2018). Malingered psychosis. In Richard Rogers & Scott D. Bender (Eds.). *Clinical Assessment of Malingering and Deception* (4th Ed., pp. 98-121). New York: Guilford.

Rogers, R. (1990). Models of feigned mental illness. *Professional Psychology: Research and Practice*, 21(3), 182–188. doi: 10.1037/0735-7028.21.3.182

Rogers, R. (2018a). An Introduction to Response Styles. In Richard Rogers & Scott D. Bender (Eds.). *Clinical Assessment of Malingering and Deception* (4th ed., pp. 3-40). New York: Guilford.

Rogers, R. (2018b). Detection Strategies for Malingering and Defensiveness. In Richard Rogers & Scott D. Bender (Eds.). *Clinical Assessment of Malingering and Deception* (4th ed., pp. 18-41). New York: Guilford.

Rogers, R. (2018c). Researching response styles. In Richard Rogers & Scott D. Bender (Eds.). *Clinical Assessment of Malingering and Deception* (4th ed., pp. 592-614). New York: Guilford.

Rogers, R., & Cruise, K. R. (1998). Assessment of malingering with simulation designs: Threats to external validity. *Law and Human Behavior*, 22(3), 273–285. doi:10.1023/a:1025702405865

Rogers, R., Gillis, J. R., Dickens, S. E., & Bagby, R. M. (1991). Standardized assessment of malingering: Validation of the Structured Interview of Reported Symptoms. *Psychological Assessment: A Journal of Consulting and Clinical Psychology*, 3(1), 89–96. doi: 10.1037/1040-3590.3.1.89

Rogers, R., Hinds, J. D., & Sewell, K. W. (1996). Feigning Psychopathology Among Adolescent Offenders: Validation of the SIRS, MMPI-A, and SIMS. *Journal of Personality Assessment*, 67(2), 244–257. doi:10.1207/s15327752jpa6702_2

Rogers, R., Harrell, E.H., & Liff, C.D. (1993). Feigning neuropsychological impairment: A critical review of methodological and clinical considerations. *Clinical Psychology Review*, 13, 255-274. doi: 10.1016/0272-7358(93)90023-F

Rogers, R., Salekin, R. T., Sewell, K. W., Goldstein, A., & Leonard, K. (1998). A comparison of forensic and nonforensic malingerers: A prototypical analysis of explanatory models. *Law and Human Behavior*, 22(4), 353–367. doi:10.1023/a:1025714808591

Rogers, R., Sewell, K. W., & Goldstein, A. (1994). Explanatory models of malingering: A prototypical analysis. *Law and Human Behavior*, 18, 543-552. doi: 10.1007/BF01499173

Rüsseler, J., Brett, A., Klaue, U., Sailer, M., & Münte, T. F. (2008). The effect of coaching on the simulated malingering of memory impairment. *BMC Neurology*, 8(1). doi:10.1186/1471-2377-8-37

Schroeder, R. W., Martin, P. K., Heinrichs, R. J., & Baade, L. E. (2019). Research methods in performance validity testing studies: Criterion grouping

approach impacts study outcomes, *The Clinical Neuropsychologist*, 33(3), 466-477, doi: 10.1080/13854046.2018.1484517

Sellbom, M., Toomey, J. A., Wygant, D. B., Kucharski, L. T., & Duncan, S. (2010). Utility of the MMPI-2-RF (Restructured Form) validity scales in detecting malingering in a criminal forensic setting: A known-groups design. *Psychological Assessment*, 22(1), 22–31. doi:10.1037/a0018222

Silk-Eglit, G. M., Lynch, J. K., & McCaffrey, R. J. (2016). Validation of Victoria Symptom Validity Test Cutoff Scores among Mild Traumatic Brain Injury Litigants Using a Known-Groups Design. *Archives of Clinical Neuropsychology*, 31(3), 231–245. doi:10.1093/arclin/acv108

Simões, M. R. (2005). O exame dos comportamentos de simulação em avaliação (neuro)psicológica. In C. Vieira, A. M. Seixas, A. Matos, M. P. Lima, & M. Vilar (Eds.), *Ensaio sobre o comportamento humano: Do diagnóstico à intervenção. Contributos nacionais e internacionais* (pp. 453-481). Coimbra: Almedina.

Simões, M. R. (2006). Testes de validade de sintomas na avaliação de comportamentos de simulação. In A. C. Fonseca, M. R. Simões, M. C. Taborda-Simões, & M.S. Pinho (Eds.), *Psicologia forense* (pp. 279-309). Coimbra: Almedina.

Simões, M. R. (2012). Simulação, esforço insuficiente e exagero de sintomas em avaliação neuropsicológica forense. In F. Almeida, & M. Paulino (Eds.), *Profiling, vitimologia e ciências forenses: Perspectivas actuais* (pp. 147-166). Lisboa: Pactor/Lidel.

Simões, M. R., Almiro, P. A., Nunes, A. P., Cunha, C., Almeida, D., Domingues, A. F., Ferreira, S., Paiva, J., Dias, I. T., & Cruz, I. (2017). Escala de Validade de Sintomas – Versão 2 (EVS-2). In M. R. Simões, L. S. Almeida, & M. M. Gonçalves (Coords.), *Psicologia Forense: Instrumentos de Avaliação* (pp. 417-435). Lisboa: Pactor.

Simões, M. R., Almiro, P. A., Mota, M., Oliveira, C., Maior, F. S., Cunha, C., Domingues, A. F., Pinho, J. I., Soares, D.M., Almeida, D., Barreto, M., & Duarte, F., & Dias, I. T. (2017). Instrumentos de Avaliação em Psicologia Forense: Contributos da investigação para a prática profissional. In M. R. Simões, L. S. Almeida, & M. M. Gonçalves (Coords.), *Psicologia Forense: Instrumentos de Avaliação* (pp. 1-17). Lisboa: Pactor.

Simões, M. R., Almiro, P. A., Mota, M., Oliveira, C., Maior, F. S.,

Cunha, C., Domingues, A. F., Pinho, J. I., Soares, D.M., Almeida, D., Barreto, M., & Duarte, F., & Dias, I. T. (2017). Inventário Estruturado de Simulação de Sintomas (SIMS). In M. R. Simões, L. S. Almeida, & M. M. Gonçalves (Coords.), *Psicologia Forense: Instrumentos de Avaliação* (pp. 397- 415). Lisboa: Pactor.

Simões, M. R., & Sousa, L. B. (2008). Avaliação neuropsicológica em contexto forense. In A. C. Fonseca (Ed.), *Psicologia e Justiça* (pp. 361-370). Coimbra: Almedina.

Simões, M. R., Sousa, L. B., Fonseca, M. S., Garcia, S. M., Pinho, J. I., Soares, D. M., Silva, D. C., Oliveira, C. S., Mota, M., Maior, F. S., Jesus, S. S., Fernandes, S. P., Faustino, J. P., Cruz, A. P., Barreto, M., Dias, I. T., Perdiz, C. P., Duro, D., Santana, I., & Pinho, M. S. (2017). Teste de Simulação de Problemas de Memória (TOMM). In M. R. Simões, L. S. Almeida, & M. M. Gonçalves (Coords.), *Psicologia Forense: Instrumentos de Avaliação* (pp. 353-375). Lisboa: Pactor.

Slick, D. J., Sherman, E. M. S., & Iverson, G. L. (1999). Diagnostic Criteria for Malingered Neurocognitive Dysfunction: Proposed Standards for Clinical Practice and Research. *The Clinical Neuropsychologist*, 13(4), 545–561. doi:10.1076/13854046(199911)13:04;1-y;ft545

Smith, G. P., & Burger, G. K. (1997). Detection of malingering: Validation of the Structured Inventory of Malingered Symptomatology (SIMS). *Journal of the American Academy of Psychiatry and the Law*, 25, 183–189.

Soares, D. M. (2013). *Deteção de comportamentos de esforço reduzido e exagero de sintomas numa amostra de reclusos do Estabelecimento Prisional de Coimbra*. (Dissertação de Mestrado). Coimbra: Faculdade de Psicologia e de Ciências da Educação da Universidade de Coimbra,

Springman, R. E., & Vandenberg, B. R. (2009). The Effects of Test-Strategy Coaching on Measures of Competency to Stand Trial. *Journal of Forensic Psychology Practice*, 9(3), 179–198. doi:10.1080/15228930902809856

Tombaugh, P. (1996). *Test of Memory Malingering: TOMM*. Toronto: Multi-Health System.

Van Dyke, S. A., Millis, S. R., Axelrod, B. N., & Hanks, R. A. (2013). Assessing Effort: Differentiating Performance and Symptom Validity.

The Clinical Neuropsychologist, 27(8), 1234–1246. doi:10.1080/13854046.2013.835447

Van Helvoort, D., Merckelbach, H., & Merten, T. (2019). The Self-Report Symptom Inventory (SRSI) is sensitive to instructed feigning, but not to genuine psychopathology in male forensic inpatients: An initial study. *The Clinical Neuropsychologist*, 1069-1082. doi:10.1080/13854046.2018.1559359

Van Impelen, A., Merckelbach, H., Jelicic, M., & Merten, T. (2014). The Structured Inventory of Malingered Symptomatology (SIMS): A systematic review and meta-analysis. *The Clinical Neuropsychologist*, 28(8), 1336-1365. doi: 10.1080/13854046.2014.984763

Vilar-López, R., Gómez-Río, M., Santiago-Ramajo, S., Rodríguez-Fernandez, A., Puente, A., & Pérez-García, M. (2008). *Malingering detection in a Spanish population with a known-groups design*. *Archives of Clinical Neuropsychology*, 23(4), 365–377. doi:10.1016/j.acn.2008.01.007

Vilar, M., Sousa, L. B., Fonseca, M. S., Lages, H., Garcia, S. M., Pinho, J. I., Soares, D. M., Oliveira, C. S., Jesus, S. S., Fernandes, S. P., Pinho, S., Barreto, M., Dias, I. T., Perdiz, C., Duro, D., Santana, I., Firmino, H., & Simões, M. R. (2017). Teste de Memória de Rey – 15 Itens (Rey 15-IMT). In M. R. Simões, L. S. Almeida, & M. M. Gonçalves (Coords.), *Psicologia Forense: Instrumentos de Avaliação* (pp. 377-396). Lisboa: Pactor.

Vitacco, J.M., & Rogers R. (2005) Assessment of malingering in correctional settings. In: C.L. Scott, J.B. Gerbasi (Eds.), *Handbook of correctional mental health* (pp. 255-276). Washington DC: American Psychiatric Publishing.

Vitacco, M., Rogers, R., Gabel, J., & Munizza, J. (2007). An evaluation of malingering screens with competency to stand Trial patients: A known-groups comparison. *Law and Human Behavior*, 31(3), 249–260. doi:10.1007/s10979-006-9062-8

Wall, T. D., Wygant, D. B., & Gallagher, R. W. (2014). Identifying Overreporting in a Correctional Setting. *Criminal Justice and Behavior*, 42(6), 610–622. doi:10.1177/0093854814556881

Walters, G. (2006). Coping with malingering and exaggeration of psychiatric symptomatology in offender populations. *American Journal of Forensic Psychology*, 24, 21–40.

Walters, G. D., Rogers, R., Berry, D. T. R., Miller, H. A., Duncan, S. A., McCusker, P. J., Payne, J. W., & Granacher, R. P., Jr. (2008). Malingering as a categorical or dimensional construct: The latent structure of feigned psychopathology as measured by the SIRS and MMPI-2. *Psychological Assessment, 20*(3), 238–247. doi:10.1037/10403590.20.3.238

Walters, G., White, T., & Greene, R. (1988). Use of the MMPI to identify malingering and exaggeration of psychiatric symptomatology in male prison inmates. *Journal of Consulting and Clinical Psychology, 56*, 111–117. doi: 10.1037/0022-006X.56.1.111

Wasyliw, O. E., Grossman, L. S., Haywood, T. W., & Cavanaugh Jr., J. L. (1988). The Detection of Malingering in Criminal Forensic Groups: MMPI Validity Scales. *Journal of Personality Assessment, 52*(2), 321–333. doi:10.1207/s15327752jpa5202_13

Widows, M. R., & Smith, G. P. (2005). *Structured Inventory of Malingered Symptomatology: Professional manual*. Odessa, FL: Psychological Assessment Resources.

Wisdom, N. M., Callahan, J. L., & Shaw, T. G. (2010). Diagnostic utility of the Structured Inventory of Malingered Symptomatology to detect malingering in a forensic sample. *Archives of Clinical Neuropsychology, 25*(2), 118-125. doi: 10.1093/arclin/acp110

Wygant, D. B., Walls, B. D., Brothers, S. L., & Berry, D. T. (2018). Assessment of Malingering and Defensiveness on the MMPI-2 and MMPI-2-RF. In Richard Rogers & Scott D. Bender (Eds.). *Clinical Assessment of Malingering and Deception* (4th Ed., pp. 257-279). New York: Guilford.

Yates, G. P., Mulla, M. M., Hamilton, J. M., & Feldman, M. D. (2018). Factitious Disorders in Medical and Psychiatric Practices. In Richard Rogers & Scott D. Bender (Eds.). *Clinical Assessment of Malingering and Deception* (4th Ed., pp. 212-235). New York: Guilford.