UNIVERSIDADE D
COIMBRA

Ana Cláudia Serdoura Venâncio

SELF-REPORT SYMPTOM INVENTORY (SRSI):
ESTUDOS DE VALIDAÇÃO EM CONTEXTO PRISIONAL (*DESIGN* DE SIMULAÇÃO ANÁLOGO, VALIDADE CONVERGENTE E VALIDADE DISCRIMINANTE)

Dissertação no âmbito do Mestrado em Psicologia, área de especialização em Psicologia Clínica e da Saúde, área de 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 2021

***Self-Report Symptom Inventory* (SRSI):**
Estudos de validação em contexto prisional (*design* de simulação análogo, validade convergente e validade discriminante)

Resumo

A adoção de estilos de resposta enviesados, como a simulação e exagero de sintomas psicopatológicos, por parte de um examinando, constitui uma ameaça à validade dos resultados de uma avaliação (neuro)psicológica, especialmente em contextos forenses, nos quais ganhos secundários têm elevada probabilidade de manifestação. Os Testes de Validade de Sintomas (TVS) têm como objetivo detetar a presença de tais estilos de resposta/ comportamentos, viabilizando, desse modo, interpretações mais fidedignas acerca do perfil de funcionamento psicológico real de um indivíduo.

A presente investigação procurou dar continuidade ao processo de validação do S*elf-Report Symptom Inventory* (SRSI; Merten et al., 2016) no contexto prisional português. Para a concretização do referido objetivo, foi administrado a uma amostra prisional de 100 reclusos, com idades compreendidas entre os 18 e os 72 anos, um protocolo de avaliação constituído pelos seguintes instrumentos: o SRSI, o *Structured Inventory of Malingered Symptomatology* (SIMS; Smith & Burger, 1997; Widows & Smith, 2005; versão portuguesa: Simões, Almiro, Mota, et al., 2017), a Escala de Validade de Sintomas – Versão 2 (EVS-2; Simões, Almiro, Nunes, et al., 2017; Cunha, 2011) e o *Brief Symptom Inventory* (BSI; Derogatis, 1982; versão portuguesa: Canavarro, 2007). O presente estudo de validação do SRSI recorreu ao *design* de simulação análogo, numa vertente intra-sujeitos. Desta forma, todos os participantes responderam ao SRSI em dois momentos contíguos: de forma honesta (condição "honesta") e de acordo com instruções (*cover story* – condição de "simulação"), elaboradas considerando recomendações da literatura científica e orientadas para maximizar a possibilidade de generalização dos dados.

No que concerne aos resultados, os estudos de fiabilidade (consistência interna das pontuações no SRSI) indicam valores elevados de *alfa de Cronbach* para ambas as escalas, em ambas as condições. No âmbito dos estudos de validade convergente, é de referir que o SRSI contém validade com outros TVS (e.g., SIMS e EVS-2), já que os resultados evidenciaram a existência de relações positivas, elevadas e significativas entre a escala de *Pseudo-sintomas* do SRSI e o resultado total do SIMS e da EVS-2 (instrumentos que medem constructos teóricos semelhantes). No mesmo sentido, e de acordo com o esperado, foram encontradas correlações positivas, moderadas e/ou fortes entre a escala de *Sintomas Genuínos* e os índices globais do BSI (instrumento que mede a presença de psicopatologia genuína).

No domínio da validade discriminante, esperavam-se correlações não significativas, reduzidas ou nulas entre a escala *Pseudo-sintomas* do SRSI e o BSI, uma vez que mensuram constructos não sobrepostos. Porém, os resultados sugeriram que a escala *Pseudo-sintomas* se revelou igualmente sensível à presença de psicopatologia, uma vez que se observaram correlações moderadas e fortes entre a mesma e os índices globais do BSI – aspeto

sinalizador de ausência de validade discriminante. Por sua vez, e de acordo com o esperado, foram observadas correlações positivas reduzidas entre o total da escala *Sintomas Genuínos* (na condição "honesta") e o total da escala *Pseudo-sintomas* (na condição "simulação").

Ainda no que concerne à validade discriminante, foi possível verificar a capacidade do SRSI, diferenciar as pontuações da condição "honesta" e da condição "simulação", na medida em que se observaram diferenças estatisticamente significativas entre as médias obtidas nas mesmas, sendo o tamanho de efeito correspondente grande. No mesmo sentido, o SRSI revelou ser capaz de discriminar as respostas dos participantes reclusos de participantes não-reclusos (amostra de controlo), pontuando os primeiros significativamente mais do que os segundos, nas escalas e na maioria das subescalas do SRSI.

São analisadas as implicações do presente estudo em contexto prisional e apresentadas recomendações para futuras investigações.

Palavras-chave: SRSI, validade de sintomas, simulação/exagero de sintomas, *design* de simulação análogo.

Self-Report Symptom Inventory (SRSI):
Validation studies in a prison context (analogous simulation *design*, convergent validity and discriminant validity)

Abstract

The adoption of biased response styles, such as the simulation and exaggeration of psychopathological symptoms, by an examinee, composes a threat to the validity of the results of a (neuro)psychological assessment, especially in forensic contexts, in which secondary gains can stand out. Symptom Validity Tests (SVT) aim to detect the presence of such response styles/ behaviors, in order to provide mental health professionals more reliable interpretations of an individual's actual psychological functioning profile.

This study seeked to continue the validation process of the S*elf-Report Symptom Inventory* (SRSI; Merten et al., 2016) for the portugueses prison context. To achieve the mentioned goal, to a prison sample of 100 inmates, aged between 18 and 72 years old, was administered an assessment protocol consisting of the following instruments: the SRSI, the *Structured Inventory of Malingered Symptomatology* (SIMS; Smith & Burger, 1997; Widows & Smith, 2005; portuguese version: Simões, Almiro, Mota, et al., 2017), the Symptom Validity Scale – Version 2 (EVS-2; Simões, Almiro, Nunes, et al., 2017; Cunha, 2011) and the *Brief Symptom Inventory* (BSI; Derogatis, 1982; portuguese version: Canavarro, 2007). The present SRSI validation study was based on analogous simulation *design*, in an intra-subject way. Therefore, all participants answered to the SRSI in two contiguous moments: in a honest way ("honest" condition) and according to instructions (*cover story* – "simulation" condition) built in line with the recommendations that emerge from the scientific literature, in an attempt to maximize the possibility of generalizability of the data.

Regarding the results, the reliability studies (internal consistency of the SRSI scores) indicate high values of *Cronbach's alpha* for both scales, in both conditions. In the context of convergent validity studies, it should be noted that the SRSI has validity with other TVS (e.g., SIMS and EVS-2), since the results showed the existence of positive, high and significant relationships between the *Pseudo-symptoms* scale (SRSI) and the total result of the SIMS and EVS-2 (instruments that measure similar theoretical constructs). In the same sense, and as expected, positive, moderate and/or strong correlations were observed between the *Genuine Symptoms* scale (SRSI) and the BSI global indices (instrument that measures the presence of genuine psychopathology).

In the domain of discriminant validity, non-significant, reduced or null correlations were expected between the *Pseudo-symptoms* scale (SRSI) and the BSI, due to the measurement of non-overlapping constructs. Nevertheless, the results suggested that *Pseudo-symptoms* scale proved to be equally sensitive to the presence of psychopathology, since there were moderate and strong correlations between it and the global BSI indices – aspect that indicates lack of discriminant validity. In turn, as expected, positive but reduced correlations were observed between the total of the *Genuine*

Symptoms scale (in the “honest” condition) and the total of the *Pseudo-symptoms* scale (in the “simulation” condition).

Also with regard to discriminant validity, it was possible to verify the SRSI's ability to differentiate the scores of the “honest” condition and the “simulation” condition, since there were statistically significant differences between the means obtained in them, with the corresponding effect size being large. In the same sense, the SRSI proved to be able to discriminate the responses of inmates from non-inmates (control sample), scoring the former significantly more than the latter, on the scales and in most subscales of the SRSI.

The implications of this study in a prison context are analyzed and recommendations for future investigations are presented.

Key Words: SRSI**, s**ymptom validity, feigning/over-reporting, analog simulation *design*.

Agradecimentos

No finalizar desta etapa, é imperioso o agradecimento àqueles que, de uma forma crucial e marcante, me auxiliaram ao longo desta caminhada:

Ao Professor Doutor Mário Simões, pela orientação, pela disponibilidade, pelo incentivo e pela excelência que o caracteriza.

À Professora Doutora Isabel Alberto, pelo suporte e pela fonte de motivação que demonstrou ser, para mim, ao longo do último ano letivo.

Aos meus pais e ao meu irmão, que me completam de uma forma inigualável e me dão a mão no final de cada dia. Por caminharem ao meu lado com a maior das certezas e por serem a minha maior força.

Aos meus amigos, por me abraçarem como parte vossa. Pela paciência e pelas palavras de coragem que, diariamente, me aconchegam. Por serem a prova de que o *amor* transcende distâncias, percursos improváveis e almas distintas.

Ao Estabelecimento Prisional da Guarda e à sua equipa técnica, por me terem recebido com a maior amabilidade e por me terem permitido *voar*. Em especial à Doutora Mónica Lopes, pela admiração e pelos conhecimentos que, atenciosamente, me transmitiu… no fundo, por ter contribuído para que o estágio curricular se traduzisse numa das experiências mais gratificantes do meu percurso académico.

Estarei, para sempre, grata.

Índice

Introdução 1

I – Enquadramento conceptual 2

1. Avaliação (neuro)psicológica em contexto forense e estilos de resposta. 2

1.1. Simulação e exagero de sintomas | Definição concetual, implicações diagnósticas e modelos explicativos. 3

2. Avaliação da validade do autorrelato | A dicotomia entre medidas integradas (*embedded*) *versus* medidas isoladas (*stand-alone*). 5

2.1. Instrumentos clássicos de avaliação (neuro) psicológica. 6

2.2. Testes de Validade de Sintomas e de Desempenhos 7

2.2.1. O *Self-Report Symptom Inventory* (SRSI). 8

3. Designs de investigação: O *design* de simulação análogo. 9

3.1. Simuladores instruídos: A relevância do *coaching*. 11

3.2. Simuladores instruídos: A relevância da *cover story*. 11

3.2.1. Questões éticas a considerar nos estudos com simuladores instruídos. 12

II - Objetivos 13

III - Metodologia 14

IV - Resultados 19

V - Discussão dos resultados 35

VI - Conclusões 44

Bibliografia 47

Anexos 57

Introdução

Os estilos de resposta e, em particular, a simulação e o exagero de sintomas constituem um tópico de investigação simultaneamente atual e relevante em Psicologia Forense. O crescente interesse por esta problemática compreende-se pela necessidade de resposta a solicitações de avaliação (neuro)psicológica, advindas de diversas conjunturas litigiosas, marcadas pela presença de incentivos externos ou ganhos secundários, frequentemente intrínsecos a contextos forenses (e.g., o acesso a medicação, indeminizações de natureza monetária, etc.; Heilbronner et al., 2009; Heilbronner & Bush, 2017). A manifestação de estilos de resposta enviesados, orientada para a obtenção de ganhos secundários, prejudica a credibilidade do (auto)relato do sujeito examinando e dos resultados da sua avaliação. Neste sentido, têm sido desenvolvidos e integrados nos protocolos de avaliação testes de validade de desempenhos (TVD) e testes de validade de sintomas (TVS), que visam determinar a credibilidade dos relatos e, desse modo, dar resposta ao potencial problema da simulação ou exagero de sintomas (Merten & Merckelbach, 2013; Merten et al., 2016; Simões, Almeida, et al., 2017).

A investigação e as diretrizes internacionais reiteram a necessidade do uso de múltiplas medidas (testes de validade) e estratégias para a deteção destes estilos de resposta (enviesamentos negativos), já que, por um lado, tal prática aumenta a probabilidade da identificação da sua presença e, por outro, reduz a possibilidade de sucesso da conduta de simulação (Larrabee, 2008). Ou seja, são necessários diversos instrumentos de avaliação da validade de desempenhos e sintomas, que permitam a deteção de condutas de simulação e de exagero de sintomatologia.

É neste contexto que se afigura a utilidade da presente investigação empírica, que dá continuidade ao processo de validação do *Self-Report Symptom Inventor*y (SRSI; Merten et al., 2016) para a população prisional portuguesa. O SRSI é um instrumento de validade de sintomas recentemente construído com o intuito de colmatar as fragilidades conhecidas associadas a outros TVS (e.g., SIMS). Neste sentido, a presente investigação inclui estudos de validade convergente, com base no estudo das relações entre medidas que avaliam constructos semelhantes, nomeadamente entre os resultados obtidos no SRSI e noutros instrumentos de avaliação administrados (e.g., SIMS, EVS-2); e de validade discriminante, através, por um lado, do estudo das relações entre as medidas que mensuram constructos distintos, e, por outro, da análise da capacidade do SRSI em diferenciar a *performance* inerente a diferentes grupos.

Adicionalmente, é de referir que a presente investigação recorre ao *design* de simulação análogo (DSA), numa vertente intra-sujeitos, paradigma muito utilizado na validação desta tipologia de instrumentos, uma vez que permite a comparação de respostas entre duas condições: uma condição de resposta "honesta" (na qual é solicitado que o participante responda com sinceridade, de forma verdadeira) e uma condição de resposta de "simulação" (na qual o participante é instruído a responder de forma a falsear ou exagerar

sintomas). A pertinência do referido *design* recai igualmente na possibilidade de estudar as respostas "honestas" ao TVS, comparando-as às respostas de um examinando previamente instruído a responder de forma simulada (conduta de *coaching*, cada vez mais prevalente na atualidade; Victor & Abeles, 2004).

I – Enquadramento conceptual

1. Avaliação (neuro)psicológica em contexto forense e estilos de resposta

A avaliação (neuro) psicológica em contexto forense objetiva traduzir evidências precisas acerca do perfil de funcionamento de um sujeito. Para tal, é imprescindível que as respostas aos diversos instrumentos que compõem o protocolo de avaliação sejam consideradas válidas, honestas e proporcionadas com base num esforço e motivação máximos (Brush et al., 2014; Dandachi-FitzGerald et al., 2020; Larrabee, 2007). Assim, é facilmente percetível que, se a avaliação não for ministrada nestas condições, esperar-se-ão resultados – que enformam o perfil psicológico obtido – potencialmente inválidos, independentemente do(s) motivo(s) inerente(s) ao estilo de resposta adotado pelo sujeito (Heide et al., 2017; Merten & Merckelbach, 2013; Rogers & Cruise, 1998).

Um estilo de resposta enviesado representa, segundo McGrath e colaboradores (2010), uma atitude ou tendência de resposta consistente a um instrumento administrado, que não traduz o verdadeiro modo de funcionamento do indivíduo. Os estilos de resposta não devem ser concetualizados como estáticos, uma vez que são multideterminados e, por isso, adequáveis aos objetivos pessoais do sujeito, específicos de cada contexto (Rogers, 2018a). De facto, a prevalência de enviesamentos de resposta varia em função do contexto no qual a avaliação está a ser administrada, pelo que examinandos inseridos numa conjuntura litigiosa apresentam um risco acrescido de os manifestar, tendo em conta a presença de potenciais e elevados ganhos secundários para que deturpem o seu verdadeiro funcionamento psicológico (e.g., procura de indeminizações monetárias; acesso a medicação, etc.). Porém, convém realçar que, ainda que o autorrelato (*self-report*) de sintomas seja, de forma notoriamente ingénua, mais comummente considerado confiável em contextos clínicos não forenses, estilos de resposta igualmente enviesados podem ser produzidos nos mesmos (Giger & Merten, 2019; Kemp et al., 2008; Martin & Schroeder, 2020; Mittenberg et al., 2002).

Nos últimos anos, destaca-se a investigação inerente aos diversos estilos de resposta, particularmente aos comportamentos de *simulação (malingering)* e de exagero *(overreporting)* de sintomas, adotados pelo examinando no contexto de avaliação (neuro)psicológica forense, no âmbito do exame da validade das suas respostas (e.g., Boone, 2021; Bush et al., 2014; Horton & Reynolds, 2021; Rogers, 2018a).

1.1. Simulação e exagero de sintomas | Definição concetual, implicações diagnósticas e modelos explicativos

Um dos estilos de resposta mais prevalente recai na tentativa de fabricar e/ou exagerar psicopatologia, forma enviesada de resposta que tem ganho foco na literatura e investigação científica (Heilbrun, 1992), pelo que será analisado na presente rubrica.

Ao longo das últimas décadas, o constructo de *simulação* tem sido alvo de debate e aperfeiçoamento, sobretudo ao nível conceptual e dos critérios de diagnóstico (Sherman et al., 2020). Atualmente, a *simulação* associa-se à produção intencional de sintomas físicos ou psicopatológicos falsos, seja pela sua fabricação total, ou pelo seu exagero, impulsionada por potenciais incentivos externos, como financeiros (e.g., indeminizações monetárias por invalidez), legais (e.g., redução da responsabilidade criminal), e outros benefícios, como a obtenção de medicação (American Psychiatric Association – APA, 2013). No *Manual de Diagnóstico e Estatística das Perturbações Mentais* (DSM-5), são enumerados três elementos subjacentes à sua definição: a produção/exagero de sintomas, a motivação e intencionalidade consciente do indivíduo para simular, e a presença de incentivos externos.

A 5ª edição do DSM (APA, 2013) não concetualiza a simulação como uma perturbação mental diagnosticável[1], incluindo-a na secção das condições que deverão ser alvo de atenção clínica, dado o impacto que a sua presença pode acarretar no diagnóstico, curso ou tratamento de uma condição psiquiátrica, devendo ser excluída previamente à determinação desta (Tracy & Rix, 2017). Em consequência, o DSM-5 salienta circunstâncias específicas nas quais a possibilidade de simulação deve ser alvo de suspeição, nomeadamente: (a) estar inserido num contexto médico-legal, devido à maior probabilidade de ganhos secundários; (b) existir discrepância entre sintomas relatados e dados objetivos da observação; (c) falta de cooperação na avaliação diagnóstica ou não adesão a tratamento, e (d) presença de um diagnóstico de Perturbação de Personalidade Antissocial (PPAS; APA, 2013).

Analisando as alíneas supracitadas, são de evidenciar limitações que lhes estão subjacentes. Primeiramente, a mera inserção num contexto médico-legal/ forense, pressupondo nele a presença de incentivos externos, não deve ser vista como suficiente, uma vez que tal conceção pode levar à sobrestimação da simulação neste cenário, e subestimação em outros, como o contexto clínico não forense (Martin & Schroeder, 2020; McCarter et al.,

[1] No passado, diversos autores enumeraram critérios para que o diagnóstico de simulação pudesse concretizar-se (e.g., Greiffenstein et al., 1994; Pankrantz & Erickson, 1990; Rogers, 1990), que foram objeto de críticas por parte de outros investigadores (e.g., Slick et al., 1999). Estes atualizaram a definição de simulação, propondo simultaneamente um conjunto de critérios diagnósticos para a Simulação de Perturbação Neurocognitiva (SPN), visando integrar, pela primeira vez, informações psicométricas e comportamentais no seu diagnóstico. Todavia, e ainda que constituam o modelo mais amplamente aceite para identificar simulação de défices cognitivos, os critérios de Slick e colaboradores foram igualmente alvo de críticas (e.g., Boone, 2007; Larrabee et al., 2007), recentemente colmatadas através de uma reconceptualização (cf. Sherman et al., 2020).

2009). No que concerne à alínea b), na avaliação da validade dos sintomas não há, muitas vezes, fontes de informação objetivas disponíveis, para além dos autorrelatos do avaliado e a apreciação do perito – curiosamente, no estudo de Rassin e colaboradores (2010), apenas 4% manifesta confiança nas suas capacidades detetivas de simulação. Por sua vez, a alusão à sugestão da não cooperação por parte da pessoa que simula (e.g., em avaliação psicológica) parece não ter em consideração, por um lado, que é incomum um indivíduo honesto estar disposto a discutir detalhes de uma experiência traumática (Resnick & Knoll, 2018), e, por outro, que simuladores tendem a apresentar as suas queixas da forma mais elaborada e detalhada possível, contrariamente ao relato breve de pessoas cujas queixas são genuínas (*ilusão da transparência*; cf. Savitsky & Gilovich, 2003). Finalmente, também a sugestão inerente à alínea d) apresenta problemas: pessoas com traços de personalidade antissocial não optam mais pela fabricação de sintomas, nem são melhores a fazê-lo (Impelen et al., 2017; Rogers & Cruise, 1998), pelo que erros inerentes ao enviesamento confirmatório que daí pode advir (considerar a simulação como consequência inevitável da PPAS) devem ser evitados (Kucharski et al., 2006). Desta forma, a presença de traços antissociais, *per si*, não deve ser usada como alerta para relatos de sintomas psicopatológicos desonestos, uma vez que estes podem ser concomitantes com patologia genuína (Rogers, 2018a).

Apesar das supracitadas diretrizes do DSM-5 aparentarem ser objetivas, o manual reflete uma desconsideração das recomendações da investigação produzida até então, sobretudo no que respeita aos critérios de diagnóstico presentes na sua 3ª edição (Berry & Nelson, 2010; Sherman et al., 2020). Desta forma, é importante analisar limitações adicionais. Num momento inicial, é de evidenciar a visão da simulação, por parte da APA (2013), como um fenómeno de "tudo-ou-nada" (categorial: respostas genuínas *versus* respostas simuladas), apesar de a simulação de sintomas ser, na verdade, um fenómeno de natureza dimensional (Rogers & Bender, 2012; 2018; Slick et al., 1999), podendo reportar-se num *continuum* (Larrabee, 2005). A prová-lo, aponta-se a existência de diferentes padrões de simulação, propostos por Ferguson (2004) e Resnick (1997): 1) a produção de sintomas não existentes, refletindo a *simulação pura* (sendo rara a sua observação; Rogers, 2018a); 2) a *simulação parcial*, podendo manifestar-se através do exagero de sintomas existentes (os défices provocados pela lesão são reportados como mais severos), ou da manutenção de sintomas outrora existentes que, tendo o indivíduo já recuperado dos mesmos, reporta a sua permanência ou agravamento; 3) a *falsa imputação de sintomas existentes* a uma causa etiologicamente não relacionada, sobretudo a um evento suscetível de compensação financeira (sendo este o cenário mais tipicamente observado).

Uma vez que nem todos os sintomas falsos têm a sua justificação em condutas de simulação, é fundamental que se excluam outras hipóteses diagnósticas, peculiarmente a Perturbação Factícia (apresenta-se como perturbação formalmente reconhecida pelo DSM-5). É de referir que ambas são conotadas pela produção intencional de sintomas, ou seja, partilham entre si a voluntariedade da produção de sintomas, sendo a motivação o constructo

diferenciador. Pressupõe-se que, na simulação, os indivíduos estão motivados por incentivos externos concretos, simulando de forma consciente, pelo que na Perturbação Factícia o ganho externo não se apura, considerando-se os motivos meramente determinados pelo ganho primário de satisfação de uma necessidade psíquica interna (APA, 2013; Impelen et al., 2017; Velsor et al., 2021). A capacidade de exclusão de diversas hipóteses diagnósticas revela-se uma competência fundamental por parte do perito, exatamente pela consequência nefasta que uma afirmação acerca de condutas de simulação num indivíduo, num determinado contexto, pode acarretar.[2] Desta forma, para além da credibilidade do relato daquele ser anulada, a negação de tratamento ou outro tipo de ajuda especializada pode ser consequência de um erro dessa natureza (Easton & Akehurst, 2011).

Em termos explicativos, várias teorias se têm focado no constructo da simulação, pelo que a discussão do conceito ganhou relevo quando Rogers (1990) documentou a existência de três modelos explicativos desta conduta. O *Modelo Patogénico*, intrinsecamente relacionado com o pensamento psicodinâmico, assenta no pressuposto de que a simulação representa uma tentativa mal-sucedida de controlar processos psicóticos e neuróticos inconscientes, que se expressam através de psicopatologia. Desta forma, o modelo identifica a psicopatologia genuína como causa inerente à simulação, prevendo que simuladores serão, eventualmente, doentes mentais genuínos. O *Modelo Criminológico*, associado à proposta do DSM-III-R (APA, 1987), pressupõe que o comportamento de simulação se verifica em indivíduos com PPAS, uma vez que fabricam doenças em contextos médico-legais e se mostram pouco cooperativos em avaliações e tratamentos (Rogers, 1990). O modelo mais aceite – o *Modelo Adaptativo* – considera a simulação um estilo de resposta que surge perante condições adversas (e.g., contextos litigiosos). Assim, de entre uma diversidade de escolhas possíveis, os indivíduos escolhem simular, quando esta opção apresenta maior probabilidade de sucesso (análise custo-benefício) no alcance de um objetivo (ganho secundário). Desta forma, Rogers (1990) defende a natureza adaptativa do comportamento simulatório, na medida em que este permite alcançar objetivos pessoais, não constituindo uma dimensão estável, mas sim uma resposta situacional (cf. igualmente McDermott et al., 2013).

2. Avaliação da validade do autorrelato | A dicotomia entre medidas integradas (*embedded*) *versus* medidas isoladas (*stand-alone*)

Sendo as condutas de simulação uma problemática comum no contexto forense e litigioso, a presença de metodologias que as avaliem demonstra-se

[2]Porém, não só a incapacidade de alcançar um diagnóstico diferencial pode acarretar erros. Atendendo às fragilidades inerentes à proposta do DSM-5, é de refletir o peso que o mesmo comporta e o consequente risco de erro de falsos positivos (classificar erroneamente o indivíduo como simulador; McDermott & Sokolov, 2009). Também a desconsideração de determinados fatores (e.g., emocionais, atencionais, entre outros; cf. Young, 2021) na interpretação dos resultados de TVS, que sinalizam simulação ou exagero de de sintomatologia, pode concorrer para erros diagnósticos.

crucial (Boskovic et al., 2021; Rogers & Cruise, 1998). Neste sentido, tal como se verificou uma evolução na definição de simulação, também os métodos utilizados para a sua deteção foram paulatinamente progredindo, com base em avanços nos conhecimentos teóricos e empíricos (Young, 2014).

Os testes que visam a avaliação da validade do autorrelato objetivam a avaliação dos já enunciados estilos enviesados de resposta, que podem assumir duas formas: a *simulação* e o *esforço reduzido* em tarefas cognitivas, ou a s*imulação* e o *exagero de sintomas* em escalas de psicopatologia. Estes estilos de resposta podem ser avaliados por medidas integradas ou isoladas (neste âmbito, por Testes de Validade de Desempenhos [TVD] e/ou por Testes de Validade de Sintomas [TVS]; Helvoort et al., 2019).

2.1. Instrumentos clássicos de avaliação (neuro) psicológica

Existem indicadores de validade incorporados em instrumentos neuropsicológicos clássicos, referidos como medidas integradas (*embedded measures*). Estas podem representar o início do processo de deteção de tendências de resposta não credíveis, num registo de rastreio, já que permitem detetar, através da análise dos seus resultados, a necessidade de uma avaliação posterior mais aprofundada (Edens et al., 2007; Young, 2014). Diversas são as vantagens apontadas à utilização desta metodologia integrada (Rogers & Bender, 2018), entre as quais: (a) o facto de servirem diversos propósitos simultaneamente, permitindo avaliar quer os défices genuínos, quer os simulados; (b) a economia do tempo de avaliação; (c) a avaliação da validade das respostas de forma contínua ao longo do exame; (d) proteger os desempenhos do *coaching* (ou de simuladores sofisticados), já que são menos transparentes acerca da intenção de avaliar a credibilidade de desempenhos.

No âmbito da avaliação de desempenhos, são frequentemente considerados instrumentos tradicionais como a *Wechsler Adult Intelligence Scale* (WAIS-III; Wechsler, 2008) – o subteste *Memória de Dígitos* tem sido o mais estudado, apresentando maior sustentação empírica para a deteção da simulação (Boone, 2007) – ou o *Trail Making Test* (TMT; Partington, 1983). Estes testes permitem assinalar a presença de potenciais comportamentos de simulação, através da análise de perfis de resultados, analisando os que não são expectáveis do ponto de vista neuropsicológico (e.g., um indivíduo acerta itens mais difíceis e falha itens mais fáceis: pontuação no *Trail B* > *Trail A*).

Relativamente à validade de sintomas, importa realçar escalas e índices incorporadas em inventários tradicionais da avaliação da personalidade, tais como o *Minnesota Multiphasic Personality Inventory* (MMPI-2; Butcher et al., 2001, como citado em Wall et al., 2015; cf. Simões & Ferraz, 2021) e o *Personality Assessment Inventory* (PAI; Morey, 2007, como citado em Merten et al., 2016; cf. Simões & Ferraz, 2021). Apesar destas medidas serem pertinentes na avaliação da validade de sintomas, sobretudo pelas inferências no âmbito da validade do autorrelato que possibilitam, convém ressalvar que esse não é o objetivo exclusivo das mesmas, pelo que as conclusões inferenciais nelas baseadas são limitadas (Simões, 2012). Adicionalmente, estes inventários não são sensíveis ao *coaching*, o que significa que, uma vez

na presença de simuladores sofisticados (e.g., instruídos a responder num determinado sentido), se denota uma diminuição da sua capacidade de detetar a simulação (Boccaccini et al., 2006).

2.2. Testes de Validade de Desempenhos e de Sintomas

Apesar das referidas medidas integradas em instrumentos psicológicos tradicionais proporcionarem um rastreio inicial da validade do autorrelato, quer no âmbito da avaliação dos desempenhos, quer da validade de sintomas, a deteção de comportamentos simulatórios não pode resumir-se às mesmas, pela limitada informação que, *per si*, disponibilizam neste âmbito. Assim, foram desenvolvidos instrumentos específicos para a avaliação destas dimensões, baseados em critérios psicométricos capazes de discernir respostas de indivíduos cujos défices ou sintomas apresentados são genuínos, daqueles que os simulam (Larrabee, 2007; Merten et al., 2013).

As medidas autónomas específicas de avaliação da validade do autorrelato – *stand-alone measures* – compreendem os testes específicos de Validade de Desempenhos (TVD), que viabilizam a determinação de um esforço insuficiente ou baixo desempenho intencional em testes neurocognitivos, e os Testes de Validade de Sintomas (TVS), que avaliam a simulação e o exagero de sintomas psicopatológicos (Merten et al., 2013; Merten et al., 2016).

Apesar de a primeira categoria de medidas destacada – TVD – não constituir o objeto da presente dissertação, é pertinente referir que visa avaliar se o sujeito examinando apresenta desempenhos consistentes com as suas aptidões reais. A título de exemplo, evidenciam-se os principais instrumentos investigados: o *Test of Memory Malingering* (TOMM; Tombaugh, 1996; como citado em Simões, Sousa et al., 2017) ou o *Rey 15 – Item Memory Test* (15-FIT; Rey, 1964, como citado em Vilar et al., 2017).

Por sua vez, os TVS são medidas que pretendem detetar o relato simulado ou exagerado de sintomas (e.g., identificam a manifestação de sintomas invulgarmente apresentados por pacientes genuínos – por vezes designados "pseudo-sintomas"), avaliando se o examinando comunica um relato preciso e consistente com a sua experiência real de sintomas (Larrabee, 2014). O *Structured Inventory of Malingered Symptomatology* (SIMS; Smith & Burger, 1997) é o TVS mais amplamente estudado e utilizado a nível internacional, dispondo, em Portugal, de diversos estudos de validação em contexto forense (cf. Simões, Almiro, Mota et al., 2017). O SIMS é um instrumento de autorrelato em formato de resposta dicotómica, constituído por 75 itens que refletem sintomas atípicos e bizarros, abrangendo uma grande variedade de sintomas psicopatológicos, como psicose, perturbações afetivas e ainda défices neurológicos. A referida medida evidencia propriedades psicométricas satisfatórias em várias amostras (Simões, Almiro, Mota et al., 2017; Smith & Burger, 1997; Vitacco et al., 2007), possuindo uma elevada capacidade de discriminação, com baixa variabilidade entre culturas (Nijdam-Jones & Rosenfeld, 2017). O seu pressuposto básico é o de que sujeitos simuladores, no intuito de se afigurarem credíveis, empregarão mais sintomas

atípicos, quando comparados com sujeitos cuja sintomatologia é genuína (Huhnt, 2013; Impelen et al., 2014; Seron, 2014).

No entanto, apesar das suas diversas características robustas (e.g., alta sensibilidade ao relato de sintomas atípicos; Boskovic et al., 2019; Helvoort et al., 2019), o SIMS tem sido objeto de críticas (Widder, 2011; para uma meta-análise, cf. Impelen et al., 2014). Uma importante limitação recai em valores de especificidade subótimos em amostras com psicopatologia severa genuína, particularmente aquando do uso de pontos de corte mais baixos, o que potencia a classificação errónea de sujeitos honestos como simuladores. Adicionalmente, é de referir que os itens do SIMS focam somente sintomatologia bizarra e improvável[3], o que o torna mais facilmente reconhecível como TVS (Merten et al., 2016; Impelen et al., 2014). Por sua vez, nos últimos anos, investigadores têm salientado a incapacidade do SIMS em diferenciar indivíduos com psicopatologia genuína daqueles que simulam, defendendo mesmo o não uso deste instrumento (Cernovsky, Mendonça, Ferrari et al., 2019; Cernovsky, Mendonça, Oyewumi et al., 2019). Estas limitações e críticas tem justificado o desenvolvimento e validação de instrumentos complementares ou alternativos, como o SRSI (objetivo do presente estudo).

A *Escala de Validade de Sintomas – Versão 2* (EVS-2; Cunha, 2011; Simões, Almiro, Nunes et al.., 2017), desenvolvida em Portugal, afigura-se como um adicional exemplo de medida, cujo objetivo se centra na avaliação de condutas de simulação e exagero de sintomas psicopatológicos, que constituem uma ameaça à validade do autorrelato. Finalmente, o *Inventory of Problems – 29* (IOP-29; Viglione et al., 2017; Viglione & Giromini, 2020) é um instrumento (TVS) recentemente desenvolvido, que, segundo Young e colaboradores (2020), tem evidenciado as propriedades psicométricas adequadas para um uso em avaliações forenses.

2.2.1. O *Self-Report Symptom Inventory* (SRSI)

As fragilidades inerentes ao SIMS criaram condições para o desenvolvimento de um novo instrumento específico de avaliação da simulação e exagero de sintomas – o *Self-Report Symptom Inventory* (SRSI; Merten et al., 2016). O SRSI é uma medida de autorrelato, constituído por 107 itens. Com um formato de resposta dicotómico, os itens são de fácil compreensão e requerem um nível de leitura equivalente ao 1º ciclo. O instrumento combina vantajosamente 2 escalas gerais de sintomas: uma escala de *Sintomas Genuínos* (reflete sintomas habitualmente relatados por sujeitos com psicopatologia real) e uma escala de *Pseudo-sintomas* (reflete sintomas atípicos e bizarros, raramente afetos a real psicopatologia). Ambas as escalas centrais incluem 5 subescalas, que se destinam à avaliação de sintomas

[3] Ao ser construído como medida de triagem para a deteção da simulação em contexto criminal, o SIMS tem como limitação o foco único em tipologias de psicopatologia potencialmente prevalentes nesse mesmo contexto (e.g., psicose), mas menos comuns em contextos de litígio civil (e.g., pedidos de recompensas monetárias por parte de trabalhadores; Merten et al, 2016; Merten et al., 2019).

psicopatológicos ditos plausíveis (e.g., sintomas depressivos, cognitivos, de dor ou ansiedade, etc.), ou de sintomas improváveis (e.g., que abrangem os domínios cognitivo, motor, etc). O racional subjacente ao SRSI reflete que sujeitos simuladores tenderão a apresentar sintomas genuínos e pseudo-sintomas de forma indiscriminada, esperando-se, por sua vez, que pacientes honestos (com sintomatologia genuína) reportem sintomas clínicos plausíveis, mas não pseudo-sintomas – ou, se os apresentarem, será em baixa frequência (Helvoort et al., 2019). Os autores do SRSI sugerem o uso de dois pontos de corte: um ponto de corte de rastreio (> 6 pseudo-sintomas) e um ponto de corte com objetivo diagnóstico (> 9 pseudo-sintomas; Merten et al., 2016).

As investigações internacionais têm evidenciado a utilidade do SRSI na deteção da simulação e do exagero de sintomas, apresentando resultados favoráveis, com uma taxa de deteção correta acima dos 80% (Boskovic et al., 2019), sugerindo que ambas as escalas globais são eficazes na discriminação de sujeitos simuladores relativamente a não simuladores (Boskovic et al., 2018; Geurten et al., 2018; Giger & Merten, 2019). Num estudo levado a cabo por Helvoort e colaboradores (2019), concluiu-se que a escala de *Pseudo-sintomas* aparenta ser imune ou insensível a psicopatologia genuína – cumprindo, assim, um dos pressupostos de eficácia inerente aos TVS (Giger & Merten, 2019). Deste modo, o SRSI tem-se evidenciado um instrumento útil no que concerne à sua utilização em contexto forense e/ou prisional (Helvoort et al., 2019).

Em Portugal, foram realizados estudos (e.g., Dwarkadas, 2018; Domingues, 2019; Pinheiro, 2019) cuja conclusão corroborou o facto de o SRSI se afigurar um instrumento simultaneamente promissor, útil e relativamente eficaz no que se propõe medir. Todavia, foram igualmente identificadas algumas limitações nos estudos portugueses com o SRSI, que justificam a sua inclusão nos protolocos de avaliação forense como instrumento complementar (Giger & Merten, 2019; Merten et al., 2016; Pinheiro, 2019), e não como instrumento alternativo. Apenso (cf. Anexo 1), consta um quadro de análise de estudos empíricos, cujos objetivos abrangeram o estudo do SRSI na população portuguesa e/ou no contexto prisional.

Em última instância, é de ressalvar as recomendações dos profissionais que defendem o uso de múltiplas medidas para a deteção do esforço insuficiente e simulação/exagero de sintomas (Impelen et al., 2014). Tal visa o aumento da probabilidade da correta identificação de estilos de resposta enviesados, assim como a diminuição da probabilidade de o indivíduo ser bem-sucedido na intenção de deturpar o seu real funcionamento, evitando, assim, o cometimento do erro Tipo I (Larrabee, 2007).

3. Designs de investigação: O *design* de simulação análogo

Um *design* investigativo estruturado é a base sobre a qual as restantes questões metodológicas de um estudo recaem, pelo que a avaliação dos estilos de resposta e a validação das medidas de validade de sintomas encontra-se dependente da sua realização. São quatro as tipologias de *design* usadas nos estudos de simulação: o *design* de simulação análogo (DSA; *Simulation*

Design); a comparação com grupos conhecidos (*Known-Groups Comparisons*); o *design* de prevalência diferencial (*Differential Prevalence Design*); e o *design* de critério parcial (*Partial Criterion Design*; Rogers, 2018a).

O DSA, base de análise da presente investigação, reflete o *design* mais comummente usado na investigação da simulação. Preconiza, geralmente, uma distribuição aleatória de cada sujeito por uma condição de controlo (na qual são solicitadas aos participantes respostas honestas) e por uma ou mais condições experimentais (nas quais a amostra é instruída a responder de forma simulada). Desta forma, permite a comparação das médias dos resultados obtidos em ambas as condições, sendo esperadas diferenças significativas entre elas. Possibilita, assim, analisar a capacidade, por parte do instrumento de avaliação em escrutínio, de discriminar respostas honestas, de respostas inautênticas. Adicionalmente, viabiliza uma comparação com resultados de grupos clínicos relevantes (i.e., pacientes genuínos[4]; Rogers, 2018a).

O DSA permite, ainda, realizar a investigação numa vertente intra-sujeitos (*within-subjects*), na qual cada participante responde ao protocolo implementado, seguindo as instruções afetas a ambas as condições ("honesta" e de "simulação"). Assim, possibilitar-se-á a comparação dos padrões de respostas a ambas as condições (Rogers & Bender, 2012; Rogers, 2018a).

A esta tipologia de *design* está inerente uma validade interna forte, na medida em que preconiza a existência de procedimentos que incluem instruções estandardizadas, condições operacionalizadas e incentivos. Por sua vez, são-lhe também apontadas limitações, particularmente ao nível da validade externa, que se revela problemática pelo facto de os participantes não se confrontarem verdadeiramente com as consequências inerentes a contextos reais, muitas vezes graves, do insucesso num determinado estilo de resposta enviesado (Rogers, 2018a). Esta limitação intitula-se tipicamente de *paradoxo das investigações de simulação* (Rogers & Cavanaugh, 1983; como citados em Gillard, 2010): é pedido aos participantes que cumpram as instruções de simulação, para que se possam estudar examinandos que, quando solicitados a responder de forma honesta, simulam. A generalização dos resultados também se vê prejudicada devido às diferenças motivacionais verificadas entre simuladores experimentais comparativamente a simuladores reais.

Dada a elevada prevalência e custos inerentes às condutas de simulação em contexto forense/ prisional, vários estudos têm objetivado precisar a eficácia e o grau de resistência dos TVS às repercussões do *coaching* (e condutas de simulação). Assim, há necessidade de os estudar em

[4] Porém, a referida prática não é recorrente, uma vez que a inclusão de grupos clínicos de comparação se pode tornar desafiante (e.g., necessidade de amostra numerosa). Esta ausência representa uma falha fundamental no DSA, já que perde, em consequência, a sua potencial relevância clínica, i.e., aos resultados obtidos estará afeta uma relevância clínica limitada (Rogers, 2018c; Rogers & Gillard, 2011). Assim, considerando as características de cada *design* (cf. Rogers, 2018b), Rogers (1997) sugere que, para uma validação bem-sucedida de medidas de validade de autorrelato, é relevante o recurso concomitante aos *design* de simulação análogo – com forte validade interna – e de comparação de grupos conhecidos – com forte validade externa –, para que se tire proveito das vantagens de ambos.

amostras constituídas por simuladores instruídos (através de um DSA), de modo a analisar o seu impacto nos examinandos, quando a eles submetidos. No entanto, neste contexto, são reduzidos os estudos que recorrem a este tipo de *design*, devido ao difícil acesso à população forense. As conclusões mais relevantes dos estudos com TVS em contexto prisional, cuja metodologia recaiu num DSA, encontram-se reunidos no Quadro 2 (cf. Anexo B).

3.1. Simuladores instruídos: A relevância do *coaching*

Sendo os contextos litigiosos e forenses deveras propícios à obtenção de ganhos secundários, são habitualmente verificáveis tentativas de alterar os resultados de testes de avaliação (neuro)psicológica, de modo a distorcer o real funcionamento do examinando. O *coaching* reflete tal conduta, consistindo tipicamente em fornecer, em contexto real, instruções que visam modificar os resultados de uma avaliação, de forma a favorecer os objetivos da pessoa instruída (Springman & Vandenberg, 2009; Victor & Abeles, 2004).

As fontes através das quais o *coaching* pode advir são diversas: através da atuação de advogados (forma mais típica de *coaching)*, que acreditam ser seu dever alertar os clientes acerca dos procedimentos da avaliação psicológica, particularmente acerca dos TVS (Victor & Abeles, 2004). A submissão reiterada a avaliações psicológicas, durante um processo judicial, pode também constituir uma fonte de *coaching* (Suhr & Gunstad, 2007). Além disso, o acesso a informações detalhadas, prontamente disponíveis na *Internet*, sobre testes, avaliações psicológicas e perturbações específicas, podem igualmente auxiliar o indivíduo a autopreparar-se para evitar a deteção da simulação de sintomas (Ruiz et al., 2002).

Existem diferentes tipos de *coaching*, que envolvem a comunicação de informações acerca: (a) dos sintomas típicos da perturbação a simular; (b) da presença de métodos de avaliação de validade de sintomas no protocolo de avaliação ao qual o indivíduo será submetido; (c) dos testes a serem aplicados, assim como de estratégias de resposta aos mesmos, para que seja evitada a deteção de respostas não credíveis ou irrealistas (Suhr & Gunstad, 2007).

3.2. Simuladores instruídos: A relevância da *cover story*

Os estudos de simulação análogos são frequentemente considerados de utilidade clínica limitada, sobretudo pelo questionável nível de motivação dos participantes: enquanto os simuladores em circunstâncias de vida real experienciam elevada motivação e tempo para preparação, os estudos análogos apresentam insuficiências nestes tópicos (Rogers & Gillard, 2011).

Na tentativa de ultrapassar tais limites, Rogers (2008; 2018c) teceu recomendações acerca da formulação das instruções (*cover stories*) apresentadas aos participantes das condições experimentais. As instruções objetivam uma envolvência ativa e o incremento de motivação interna dos participantes. Identificam-se seis elementos a considerar na sua formulação: devem ser (1) *compreensíveis*, de modo a que os participantes percebam o que lhes é pedido e o que é esperado de si; devem ser (2) *suficientemente*

específicas, de modo a que o investigador consiga interpretar e inferir conclusões acerca dos resultados. Relativamente aos (3) *cenários* apresentados, deve ser considerada a sua *adequabilidade* perante os participantes, na medida em que, se forem percecionados pelos mesmos como relevantes, o seu desempenho na tarefa ver-se-á, potencialmente, maximizado. Assim, revela-se pertinente promover, junto do participante, a compreensão da (4) *relevância* que o estudo pode contrair, na medida em que aquele pode admitir-se mais motivado, caso o propósito da investigação lhe pareça pertinente. É também importante estimular a (5) *motivação interna* dos participantes, para que se verifique um envolvimento genuíno na tarefa. Neste sentido, podem ser oferecidos incentivos positivos (e.g., dinheiro) ou negativos[5]. Por fim, a (6) *credibilidade* é igualmente um elemento relevante, já que a sua perda está potencialmente associada ao fracasso em atingir o objetivo das instruções (i.e., simular). Num *design* desta natureza, para que se esperem respostas tecidas da forma mais credível possível, por parte dos simuladores instruídos, é fundamental que sejam introduzidos uma sugestão e um aviso acerca da existência de indicadores de simulação nos instrumentos protocolares (monitorização da credibilidade).

Como forma de verificar o empenho depositado na tarefa, por parte dos participantes, Rogers (2018c) defende a necessidade de uma verificação da manipulação inerente às instruções experimentais (i.e., uma avaliação do empenho do participante em dar resposta ao requerido). Pode ser obtida através das seguintes estratégias: confirmação de recordação e de compreensão das instruções, e (auto)avaliação do esforço reportado.

3.2.1. Questões éticas a considerar nos estudos com simuladores instruídos

Dilemas de natureza ética emergem aquando da realização de investigações assentes num *design* de simulação, tais como a necessidade de proteção da integridade dos instrumentos (e.g., não revelar informações acerca da função do próprio TVS; Victor & Abeles, 2014). Neste contexto, é típico o recrutamento de estudantes universitários para esta tipologia de investigação (e.g., Boskovic et al., 2019; Domingues, 2019), de modo a diminuir o risco de revelação de informações cruciais a indivíduos que se encontrem em contextos litigiosos ou forenses reais.

O já enunciado paradoxo inerente às investigações de simulação –

[5] Através de um estudo realizado por Rogers e Cruise (1998), foi possível perceber que a existência de diferenças metodológicas ao nível das instruções apresentadas provocam disparidades nos resultados obtidos, particularmente quando estas diferenças recaem no incentivo fornecido. Neste estudo, incentivos negativos, em contraponto aos positivos, originaram um maior foco no relato dos sintomas, por parte dos participantes (Rogers & Cruise, 1998). Todavia, num estudo levado a cabo por Boskovic (2019), em que diferentes condições foram alvo de distintas tipologias de incentivos (positivos e negativos), os grupos não diferiram quanto à motivação para simular ou à estratégia selecionada para o fazer – contrapondo resultados de Peace e Richards (2014). Uma possível explicação para tais resultados recai numa *credibilidade* significativamente mais baixa, no que ao incentivo positivo concerne.

evidentemente presente no *design* de base deste estudo – também acarreta limitações éticas, na medida em que é pedido ao participante que aja inversamente às instruções que lhe seriam fornecidas em situações verídicas (num contexto de avaliação psicológica; Gillard, 2010). Adicionalmente, um dilema surge aquando da publicação dos estudos e inerentes resultados: para que estes possam orientar decisões clínicas, ou ser reproduzidos por outros investigadores, há necessidade de detalhar as metodologias utilizadas (e.g., estratégias de *coaching*). Tal cenário pode prejudicar a integridade do teste em causa, uma vez que as publicações são disseminadas, por vezes, em plataformas de acesso livre, o que pode resultar num uso inapropriado ou mal-intencionado das informações (Berry et al., 1994).

II - Objetivos

A presente investigação empírica tem como objetivo principal prosseguir o processo de validação do *Self-Report Symptom Inventory* (SRSI), para o contexto prisional português, previamente iniciado por Pinheiro (2019). Os seus objetivos específicos são:

1. O exame da fiabilidade das pontuações obtidas no SRSI e nos restantes testes psicológicos que integram o protocolo de avaliação (seguidamente definido), através da análise da consistência interna;
2. A realização de estudos de validade entre as pontuações obtidas no SRSI e nos restantes instrumentos psicológicos que integram o protocolo de avaliação, tendo em conta o uso de um *design* de simulação análogo para o SRSI;
3. A definição de dados descritivos, de tendência central e de dispersão, a fim de comparar as médias obtidas inerentes à amostra recolhida em contexto prisional, com aquelas que foram alcançadas noutras investigações com o SRSI, em contexto prisional: o estudo de Pinheiro (2019) e o estudo alemão de Huhnt (2013).
4. A frequência das respostas obtidas às questões que integram o Questionário Pós-experimental de verificação de manipulação.

Comparativamente ao estudo de Pinheiro (2019), a presente investigação introduz um *design* de simulação análogo (tipicamente usado para a validação de instrumentos desta natureza), assim como acrescenta o estudo da validade discriminante do SRSI. O presente estudo pretende ainda replicar alguns dos tópicos estudados na investigação de Pinheiro (2019), tais como o estabelecimento de dados descritivos, de tendência central e de dispersão (para que se potencie a suprarreferida comparação de resultados); a realização de estudos de validade com os restantes instrumentos incluídos no protocolo de avaliação, sobretudo com outros TVS (SIMS e EVS-2) e uma medida geral de psicopatologia (BSI). Desta forma, a atual investigação concretiza o primeiro estudo português de validação do SRSI com base num *design* de simulação análogo.

As hipóteses de estudo são as seguintes:

- **Hipótese 1:** São esperados valores mais elevados no SRSI, na condição de cenário de "simulação", comparativamente à condição de resposta "honesta".
- **Hipótese 2:** No domínio da validade convergente, são esperadas correlações positivas elevadas entre os resultados na escala *Pseudo-sintomas* do SRSI (instrumento a validar) e no SIMS e na EVS-2 (instrumentos já validados em contexto prisional) – medidas que avaliam constructos semelhantes no contexto do exame da simulação ou exagero de sintomatologia;
- **Hipótese 3:** Ainda no âmbito da validade convergente, são expectáveis relações positivas elevadas entre os resultados na escala *Sintomas Genuínos* do SRSI e resultados globais no BSI;
- **Hipótese 4:** No contexto da validade discriminante, são esperadas relações positivas mais reduzidas entre entre os resultados das escalas/instrumentos que avaliam constructos distintos
- **Hipótese 5:** No âmbito da validade discriminante, são esperadas diferenças estatisticamente significativas entre as pontuações obtidas em ambas as condições do SRSI ("honesta" *versus* de "simulação").
- **Hipótese 6:** Ainda no contexto da validade discriminante, são esperadas diferenças estatisticamente significativas entre as pontuações obtidas pelos participantes reclusos *versus* não-reclusos (amostra de controlo).

Os dados empíricos decorrentes da testagem das referidas hipóteses e objetivos contribuirão para concluir se o SRSI é um instrumento que pode ser incluído num protocolo de avaliação (neuro)psicológica em contexto forense (prisional) de forma complementar, ou em alternativa, a outros TVS previamente estudados e validados na população prisional portuguesa (e.g., SIMS e EVS-2).

III - Metodologia

Procedimentos

No intuito de tornar possível a realização do presente estudo em contexto prisional, houve necessidade de solicitar a sua acreditação à Direção do Estabelecimento Prisional da Guarda. Após a sua autorização, a recolha dos dados amostrais concretizou-se entre o dia 6 de abril e 2 de junho, do ano de 2021. Neste âmbito, procedeu-se a uma seleção preliminar dos sujeitos-participantes, na qual foram considerados como critérios de exclusão o analfabetismo e a existência de dificuldades na compreensão da língua portuguesa. A administração do protocolo de avaliação decorreu em gabinetes destinados ao atendimento individual de reclusos, selecionados de acordo com a disponibilidade do espaço reservado para o efeito. Durante o referido

período, foi diariamente comunicada uma listagem que integrava os nomes dos reclusos-participantes ao corpo da guarda prisional, que os conduzia ao gabinete previamente selecionado.

De forma a proporcionar condições de administração estandardizadas, a aplicação do protocolo de avaliação foi feita individualmente. Os participantes foram elucidados, de forma escrita e verbal, acerca da finalidade e do caráter voluntário da investigação, tendo-lhes sido garantido o anonimato e a confidencialidade dos dados recolhidos (cf. Anexo C). Posteriormente, foram esclarecidas eventuais dúvidas existentes.

No que concerne ao protocolo de avaliação, foram aplicados os instrumentos SRSI, BSI, SIMS e/ou EVS-2, pela ordem descrita. Cada um dos dois últimos instrumentos foi aplicado, de forma intercalada, a cada metade da amostra, decisão que se baseou na necessidade de diminuir o tempo dispensado para a administração do protoloco de avaliação, cujo preenchimento perdurou por, aproximadamente, sessenta minutos. Neste sentido, é de ressalvar o *design* de investigação inerente ao presente estudo empírico: o DSA (intra-sujeitos), que proporcionou a administração duplicada do SRSI. Num primeiro momento, foi requerido aos participantes que respondessem ao referido instrumento de forma honesta (condição "honesta"), pelo que, posteriormente ao término da aplicação dos restantes instrumentos (e já depois da apresentação e análise conjunta do cenário de simulação), foi introduzido, novamente, o SRSI (condição "simulação").

Após finalizada a recolha dos dados amostrais, procedeu-se à cotação e tratamento estatístico daqueles. Para tal, recorremos ao programa *Statistical Package for the Social Sciences* (SPSS) versão 22.0 para *Windows*.

Participantes

A concretização da presente investigação empírica contou com a oportunidade de recolher uma amostra prisional, designadamente no Estabelecimento Prisional da Guarda, composta por 100 reclusos do sexo masculino, cujas idades se compreendem entre os 18 e os 72 anos (M= 40.51; DP= 12.79). Na sua maioria, os sujeitos que integram a presente amostra são de nacionalidade portuguesa (n= 94; 94%), verificando-se igualmente a participação de indivíduos de nacionalidade angolana (n=1), francesa (n=2), venezuelana (n=1), luxemburguesa (n=1), cabo-verdiana (n=1) e espanhola (n=1).

No que concerne ao estado civil, 58 indivíduos são solteiros (58%), 21 encontram-se casados, ou em união de facto, e 21 estão divorciados. Relativamente ao nível de escolaridade concluído aquando da participação na presente investigação (Tabela 1), 39 dos participantes completaram 9 anos de ensino, 16 completaram 6 anos de ensino, 15 concluíram 12 anos e 9 sujeitos completaram 4 anos, existindo também indivíduos com 13, 16 ou 19 anos de ensino completos. Por sua vez, a maioria dos participantes (95%) nega a existência de um diagnóstico clínico na área da saúde mental, pelo que 5 indivíduos comunicam existência de um diagnóstico (e.g., perturbações do foro depressivo e ansioso).

Tabela 1. Características sociodemográficas da amostra

	n	*%*		*n*	*%*
Nacionalidade			**Anos de Escolaridade**		
Portuguesa	94	94	4 anos	9	9
Francesa	2	2	5 anos	4	4
Venezuelana	1	1	6 anos	16	16
Cabo-verdiana	1	1	7 anos	3	3
Espanhola	1	1	9 anos	39	39
Luxemburguesa	1	1	10 anos	5	5
Estado civil			11 anos	4	4
Solteiro	58	58	12 anos	15	15
Casado/ união de facto	21	21	13 anos	1	1
Divorciado	21	21	16 anos	3	3
			19 anos	1	1

Diagnóstico clínico (saúde mental)			**Idade**			
			M	**DP**	**Min.**	**Máx**
Não	95	95	40.51	12.75	18	72
Sim	5	5				

Instrumentos e/ou materiais utilizados

1. Questionário de Dados Sociodemográficos

O referido questionário (cf. Anexo D) procurou recolher informações sociodemográficas, relativas à idade, nacionalidade, estado civil e habilitações literárias (nomeadamente anos de escolaridade completos), assim como acerca da existência de diagnósticos clínicos na área da saúde mental.

2. ***Self-Report Symptom Inventory*** (SRSI; Merten et al., 2016)

O SRSI é um instrumento de autorrelato, que visa avaliar condutas de simulação e/ou de exagero de sintomas, combinando, para o efeito, cinco subescalas clínicas de *Sintomas Genuínos* e cinco subescalas clínicas de *Pseudo-sintomas*. O instrumento é composto por um total de 107 itens, dos quais 100 itens se distribuem por ambas as escalas e inerentes subescalas (10 itens cada). O SRSI inclui, ainda, um conjunto de 5 itens que se destinam à mensuração da consistência das respostas, bem como 2 itens que analisam a cooperação do examinando. O formato de resposta recai numa escala

dicotómica (verdadeiro ou falso), sendo os seus itens simples e adequados a um nível de leitura correspondente ao 1º ciclo (Merten et al., 2016).

O racional subjacente ao SRSI é o de que sujeitos honestos tenderão a apresentar mais *Sintomas Genuínos* do que *Pseudo-sintomas* (Boskovic et al., 2018). A literatura internacional apresenta, para o total da escala *Sintomas Genuínos* e *Pseudo-sintomas*, valores de *alfa de Cronbach* de .95 e .92, respetivamente (DeVellis, 2013; Giger & Merten, 2018; Merten et al., 2016). Em contexto nacional, a consistência interna para ambas as escalas globais do SRSI varia entre valores respeitáveis e muito bons (Domingues, 2019; Dwarkadas, 2018; Pinheiro, 2019). Desta forma, com base nas investigações empíricas realizadas, é de referir que o SRSI tem boas características psicométricas (fiabilidade – e.g., consistência interna, estabilidade temporal teste-reteste; e validade – e.g., validade concorrente com o SIMS), legitimando a continuação do processo de adaptação e validação do instrumento para a população portuguesa (prisional, no caso do presente estudo).

3. ***Brief Symptom Inventory*** (BSI; Derogatis, 1982; versão portuguesa: Canavarro, 2007)

O Inventário de Sintomas Psicopatológicos constitui a adaptação para a população portuguesa do *Brief Symptom Inventory* (BSI). É um instrumento de autorrelato composto por 53 itens, que permite avaliar o mal-estar sintomático ao longo de um *continuum* (desde o mal-estar psicológico com pouco ou nenhum significado clínico, ao mal-estar mórbido). Assim, pode ser aplicado não apenas a indivíduos com perturbações psiquiátricas ou psicológicas, mas à população em geral. Cada um dos seus itens é respondido com base numa escala de *Likert* de 5 pontos, em função da classificação (atribuída pelo examinando) do grau em que o problema indicado no item o afetou durante a última semana. Integra nove dimensões de sintomatologia psicopatológica e três Índices Globais (Índice Geral de Sintomas, Total de Sintomas Positivos e Índice de Sintomas Positivos), que constituem avaliações sumárias de perturbação emocional.

Na versão portuguesa, a consistência interna entre as subescalas recai entre valores indesejáveis a respeitáveis (.62 a .80; DeVellis, 2013). Por sua vez, a média das pontuações na população geral foi de 0.84 ± 0.48, pelo que na população clínica foi de 1.43 ± 0.71. O ponto de corte da versão portuguesa foi considerado um valor igual ou superior a 1.7, o que indica gravidade sintomatológica (Canavarro, 2007).

4. ***Structured Inventory of Malingered Symptomatology*** (SIMS; Smith & Burger,1997; Widows & Smith, 2005; versão portuguesa: Simões, Almiro, Mota et al., 2017).

O SIMS é uma medida de autorrelato desenvolvida como ferramenta de triagem para a deteção de comportamentos de simulação de sintomas psiquiátricos, assumindo-se como medida da validade de sintomas. É composto por 75 itens de formato de resposta dicotómica (verdadeiro ou

falso), agrupados em cinco escalas, cada uma composta por 15 itens, que refletem sintomas atípicos e bizarros inerentes a condições como Psicose (P), Défice Neurológico (NI), Perturbações Mnésicas (AM), Capacidade Intelectual Reduzida (LI) e Perturbações Afetivas (AF).

O resultado total do SIMS permite estabelecer uma suspeita de simulação. Por sua vez, a análise qualitativa das pontuações das suas escalas permite identificar a tentativa de simulação de uma perturbação clínica em particular, através da identificação dos sintomas que se apresentam de forma atípica. De acordo com Widows e Smith (2005), uma pontuação total superior a 14 pontos requer uma avaliação posterior, mais compreensiva, a fim de verificar a presença de simulação. Todavia, outros valores de pontos de corte para o SIMS têm sido sugeridos, acarretando incerteza acerca do ponto de corte a utilizar (Impelen et al., 2014; cf. igualmente Rogers et al., 2014; Wisdom et al., 2010).

O manual do instrumento (Widows & Smith, 2005) comunica uma boa fiabilidade temporal teste-reteste para a pontuação total do SIMS, assim como um bom valor de consistência interna em todas as subescalas e no resultado total, com coeficientes de *alfa* a variar entre .82 e .88. No panorama nacional, os valores de consistência interna para o total do SIMS, alcançados por meio de alguns estudos em contexto prisional e médico-legal, variam entre .81 e .91 (DeVellis, 2013; Simões, Almito, Mota et al., 2017).

5. **Escala de Validade de Sintomas – Versão 2** (EVS-2; versão final de Simões et al., 2017; Cunha, 2011)

A EVS-2 é uma medida breve (de autorrelato) de rastreio para examinar condutas de simulação e de exagero de sintomas, que integra 48 itens na sua composição, que se encontram distribuídos por 3 domínios – Psicose (P), Perturbações Emocionais (PE) e Problemas Cognitivos (PC) –, cuja tipologia de resposta recai num formato dicotómico. Originalmente desenvolvida em Portugal, a EVS-2 visa tornar mais rigorosa a avaliação dos comportamentos de simulação e de exagero de sintomas, através da avaliação de sintomas psicopatológicos, neurocognitivos e somáticos. Pretende constituir-se como uma alternativa ou complemento a instrumentos já utilizados âmbito do teste da validade de sintomas (e.g., SIMS; Simões, Almiro, Nunes et al., 2017).

Diversas foram as investigações que objetivaram o estudo da EVS-2. Numa amostra da comunidade, verificaram-se valores de consistência interna respeitáveis para o resultado total (.70) e valores inaceitáveis para as suas dimensões PC, PE e P (.33, .40 e .58, respetivamente; Simões, Almiro, Nunes et al., 2017).

6. Cenário de simulação e Questionário Pós-experimental

O cenário de simulação (*cover story*) apresentado aos reclusos-participantes pode ser analisado no Anexo E. Foi elaborado e construído de forma a refletir e integrar as recomendações tecidas por Rogers (2018c) acerca das instruções fornecidas aos participantes de uma investigação desta natureza.

O Questionário Pós-experimental (doravante QPE; cf. Anexo F) é constituído por cinco questões. A primeira questão – "*Considera o cenário compreensível (*compreende o que está a ser pedido*)*?" – foi administrada imediatamente após a leitura da *cover story* (e previamente à administração do SRSI – condição "simulação"), visando confirmar a perceção, por parte do participante, do requerido pela *cover story*. Por sua vez, as restantes questões que compõem o QPE foram respondidas no final do processo de resposta ao SRSI, na condição "simulação". A administração destas visou, por um lado, a obtenção de informações acerca do modo com que o participante se posicionou face às instruções (*cover story*), potenciando uma compreensão acerca da adequabilidade e especificidade destas (e.g., "*Considera este cenário viável e realista?*", "*Achou compreensível que o personagem simulasse?*" e "*Foi fácil para si interpretar e imaginar ser o personagem?*"); e, por outro, permitiu ir ao encontro das recomendações divulgadas pela literatura científica, acerca dos procedimentos de verificação da manipulação (Rogers, 2018c). No âmbito do último tópico, as questões 1 ("*Considera este cenário compreensível?*") e 5 ("*Fez o seu melhor para comunicar os sintomas de modo convincente e credível?*") visaram proceder a uma verificação das condições de manipulação, recorrendo às seguintes estratégias: certificação de que o participante compreendeu as instruções, confirmação de recordação (questão 1) e (auto)avaliação do esforço investido (questão 5).

IV - Resultados

Seguidamente, são apresentadas as medidas descritivas das pontuações obtidas no SRSI, assim como os resultados relativos aos estudos de fiabilidade (das pontuações obtidas nos diversos instrumentos de avaliação aplicados à presente amostra prisional) e de validade (convergente e discriminante), que suportaram o exame das Hipóteses em estudo.

1. Fiabilidade das pontuações obtidas nos instrumentos psicológicos utilizados

No que concerne ao estudo de precisão dos suprarreferidos instrumentos de avaliação, foi calculado o índice de consistência interna, recorrendo à análise estatística do coeficiente *alfa de Cronbach*, para as escalas, subescalas e respetivos itens, conforme representado nos seguintes subtópicos e inerentes tabelas. Seguimos o critério definido por DeVellis (2003), onde um valor *alfa de Cronbach* inferior a .60 é considerado inaceitável; um valor entre .60 a .65 considera-se indesejável; entre .65 a .70

é minimamente aceitável; "respeitável" é um valor entre .70 a .80; entre .80 a .90 é muito bom. É de salientar que valores superiores a .90 podem ser indicativos de uma homogeneidade "exagerada" dos itens do instrumento.

a. Fiabilidade das pontuações obtidas no SRSI, em ambas as condições administradas, através da análise da consistência interna das subescalas e escalas

O valor de *alfa* obtido para o total de ambas as escalas de *Sintomas Genuínos* e *Pseudo-sintomas* do SRSI, na condição "honesta", foi superior a .90 (.94 e .93, respetivamente) – cf. Tabela 2. No que respeita aos valores de *alfa* inerentes às respostas obtidas nas subescalas, é de referir valores indicativos de uma consistência interna muito boa para as subescalas *Sintomatologia Cognitiva* (.87); *Sintomatologia de Dor* (.87) e *Sintomatologia Somática* (.83). As restantes subescalas inerentes à escala *Sintomas Genuínos* revelaram uma consistência interna respeitável (*Sintomatologia Ansiosa*: .76) e minimamente aceitável (*Sintomatologia Depressiva*: .68).

Por conseguinte, obteve-se um valor indicativo de consistência interna muito boa dos itens que compõe a subescala *Pseudo-sintomas Cognitivos* (.83), e respeitável dos itens inerentes às subescalas *Pseudo-sintomas Sensoriais* (.73), *Pseudo-sintomas de Dor* (.74) e *Pseudo-sintomas de Ansiedade/ Depressão* (.77). Contudo, a subescala *Pseudo-sintomas Motores* evidenciou um valor de *alfa* inaceitável (.59).

Ainda, é de referir que os itens que medem a consistência interna de respostas registaram um valor de alfa estimado como respeitável (.74).

Tabela 2. Valores estimados do *Alfa de Cronbach* para as subescalas, escalas e itens que medem a consistência interna das respostas ao SRSI, na condição "honesta"

SRSI	Nº de Itens	α de Cronbach
Subescalas Sintomas Genuínos		
Sintomatologia Cognitiva	10	.867
Sintomatologia Depressiva	10	.684
Sintomatologia de Dor	10	.874
Sintomatologia Somática	10	.826
Sintomatologia Ansiosa	10	.764
Total Sintomas Genuínos	50	.944
Subescalas Pseudo-Sintomas		
Pseudo-Sintomas Cognitivos	10	.826
Pseudo-Sintomas Motores	10	.586
Pseudo-Sintomas Sensoriais	10	.734
Pseudo-Sintomas de Dor	10	.736
Pseudo-Sintomas de Ansiedade/Depressão	10	.772
Total Pseudo-Sintomas	50	.931
Consistência de Resposta	5	.737

Em relação ao valor de consistência interna obtido para o total de ambas as escalas de *Sintomas Genuínos* e *Pseudo-sintomas*, do SRSI, na condição "simulação" (cf. Tabela 3), são igualmente de referir valores de *alfa* superiores a .90 (.94 e .98, respetivamente). No que concerne aos valores de *alfa* inerentes às respostas obtidas nas subescalas, evidenciam-se valores indicativos de uma consistência superior a .90 (subescala *Sintomatologia de Dor*) e muito boa para as subescalas *Sintomatologia Cognitiva* (.84) e *Sintomatologia Somática* (.83). As restantes subescalas da escala *Sintomas Genuínos* revelaram uma consistência interna respeitável (*Sintomatologia Ansiosa:* .73) e minimamente aceitável (*Sintomatologia Depressiva*: .66).

Por sua vez, obtiveram-se valores de consistência interna muito bons para todas as subescalas que compõe a escala *Pseudo-sintomas*, com valores a oscilar entre .82 (subescala *Pseudo-sintomas de Dor*) e .89 (*Pseudo-sintomas Motores*). É ainda de referir que os itens que medem a consistência interna de respostas registaram um valor de *alfa de Cronbach* estimado como minimamente aceitável (.697).

Tabela 3. Valores estimados do *Alfa de Cronbach* para as subescalas, escalas e itens que medem a consistência interna das respostas ao SRSI, na condição "simulação"

SRSI	Nº de Itens	α de Cronbach
Subescalas Sintomas Genuínos		
Sintomatologia Cognitiva	10	.841
Sintomatologia Depressiva	10	.657
Sintomatologia de Dor	10	.909
Sintomatologia Somática	10	.827
Sintomatologia Ansiosa	10	.729
Total Sintomas Genuínos	50	.938
Subescalas Pseudo-Sintomas		
Pseudo-Sintomas Cognitivos	10	.870
Pseudo-Sintomas Motores	10	.892
Pseudo-Sintomas Sensoriais	10	.887
Pseudo-Sintomas de Dor	10	.820
Pseudo-Sintomas de Ansiedade/Depressão	10	.851
Total Pseudo-Sintomas	50	.984
Consistência de Resposta	5	.697

b. Fiabilidade das pontuações obtidas no SIMS, na EVS-2 e no BSI, através da análise da consistência interna das suas escalas

Os valores de *alfa de Cronbach* relativos às respostas ao SIMS (cf. Tabela 4) refletem valores inaceitáveis de .54, .29 e .05, para as escalas *Psicose* (P), *Défice Neurológico* (NI) e *Capacidade Intelectual Reduzida* (LI),

respetivamente. Por sua vez, a escala *Perturbações Afetivas* (AF) obteve um valor indesejável (.64) e a escala *Perturbações Mnésicas* (AM) um valor minimamente aceitável (.67).

No que respeita ao Total do SIMS (para 75 itens), é de evidenciar um valor de consistência interna muito bom (.83).

Tabela 4. Valores estimados do *Alfa de Cronbach* para as escalas do SIMS

SIMS	Nº de Itens	α de Cronbach
Escala NI (Défice Neurológico)	15	.295
Escala AF (Perturbações Afetivas)	15	.645
Escala P (Psicose)	15	.537
Escala LI (Capacidade Intelectual Reduzida)	15	.052
Escala AM (Perturbações Mnésicas)	15	.665
Total	75	.825

Ao nível da consistência interna, a EVS-2 obteve valores de *alfa* inaceitáveis (*Escala Problemas Cognitivos*: .50), minimamente aceitáveis (*Escala Perturbações Emocionais*: .66) e respeitáveis (*Escala Psicose*: .75). Não obstante, na Tabela 5, pode ainda verificar-se um valor de consistência interna muito bom (.86) para a pontuação Total da EVS-2.

Tabela 5. Valores estimados do *Alfa de Cronbach* para as escalas da EVS-2

EVS-2	Nº de Itens	α de Cronbach
Escala PE (Perturbações Emocionais)	14	.656
Escala P (Psicose)	20	.750
Escala PC (Problemas Cognitivos)	14	.503
Total	48	.858

A análise dos valores de *alfa de Cronbach* para o BSI (cf. Tabela 6) mostra que, para as dimensões *Obsessões-Compulsões*, *Depressão*, *Hostilidade e Ansiedade*, os valores obtidos de .80, .82, .84 e .86, respetivamente, são indicativos de uma consistência interna muito boa. Três das restantes dimensões do BSI evidenciaram um valor de *alfa* respeitável: *Somatização* (.79), *Sensibilidade Interpessoal* (.77) e *Ideação Paranóide* (.75). Já ambas as dimensões *Ansiedade Fóbica* e *Psicoticismo* revelaram uma consistência interna indesejável (.62 e .63, respetivamente).

Tabela 6. Valores estimados do *Alfa de Cronbach* para as escalas do BSI

BSI	Nº de Itens	α de Cronbach
Dimensão I: Somatização	7	.793
Dimensão II: Obsessões-Compulsões	6	.804
Dimensão III: Sensibilidade Interpessoal	4	.765
Dimensão IV: Depressão	6	.816
Dimensão V: Ansiedade	6	.856
Dimensão VI: Hostilidade	5	.836
Dimensão VII: Ansiedade Fóbica	5	.618
Dimensão VIII: Ideação Paranóide	5	.751
Dimensão IX: Psicoticismo	5	.633

2. Medidas descritivas das pontuações obtidas no SRSI em ambas as condições experimentais

Na Tabela 7, apresenta-se a análise descritiva dos resultados obtidos no SRSI (escalas e subescalas), designadamente na condição "honesta", onde se encontram as pontuações médias obtidas (incluindo as pontuações mínimas e máximas – amplitude) e os respetivos desvios-padrão.

A escala *Pseudo-sintomas* apresenta uma pontuação média mais baixa, quando comparada com a escala *Sintomas Genuínos*, cujas médias são, respetivamente, de 3.62 (*DP*= 6.10) e de 15.30 (*DP*= 10.92), de acordo com o que seria de esperar. No que respeita à escala *Sintomas Genuínos*, a subescala com pontuação média mais elevada foi *Sintomatologia Ansiosa* (*M*= 4.48, *DP*=2.50), em contraponto à subescala *Sintomatologia de Dor*, que registou o valor médio mais baixo (*M*=2.34, *DP*=2.77). Por sua vez, nas subescalas pertencentes à escala *Pseudo-sintomas*, o valor médio mais baixo respeita à subescala *Pseudo-sintomas Motores* (*M*=.56, *DP*=1.06) e o valor mais elevado registou-se na subescala *Pseudo-sintomas de Ansiedade/Depressão* (*M*=.99, *DP*=1.67).

Tabela 7. Análise descritiva (médias, desvios-padrão e amplitude) dos resultados no SRSI na condição "honesta"

SRSI	M	DP	Min.	Máx.
Subescalas Sintomas Genuínos				
Sintomatologia Cognitiva	2.46	2.79	0	10
Sintomatologia Depressiva	3.23	2.06	0	9
Sintomatologia de Dor	2.34	2.77	0	10
Sintomatologia Somática	2.79	2.77	0	10
Sintomatologia Ansiosa	4.48	2.50	0	10
Total Sintomas Genuínos	15.30	10.92	1	42
Subescalas Pseudo-Sintomas				
Pseudo-Sintomas Cognitivos	.79	1.66	0	7
Pseudo-Sintomas Motores	.56	1.06	0	4
Pseudo-Sintomas Sensoriais	.63	1.30	0	9
Pseudo-Sintomas de Dor	.65	1.34	0	7
Pseudo-Sintomas de Ansiedade/Depressão	.99	1.67	0	9
Total Pseudo-Sintomas	3.62	6.10	0	29

Na Tabela 8, destacam-se as pontuações médias obtidas no âmbito da análise descritiva dos resultados no SRSI (escalas e subescalas), na condição "simulação". Relativamente às subescalas da escala geral *Sintomas Genuínos*, a pontuação média mais elevada registada foi de 8.68 (*DP*=2.11), sendo correspondente à subescala *Sintomatologia Somática*, e a mais baixa refere-se à subescala *Sintomatologia Depressiva* (*M*=7.24, *DP*=2.16). Por sua vez, nas subescalas da escala *Pseudo-sintomas* do SRSI, a média mais alta verificou-se na subescala *Pseudo-sintomas de Ansiedade/ Depressão* (*M*=6.25, *DP*=3.06), e a mais baixa na subescala *Pseudo-sintomas Sensoriais* (*M*=4.47, *DP*=3.45).

Em relação às escalas globais do SRSI, a escala de *Sintomas Genuíno*s e a de *Pseudo-sintomas*, a média mais alta registou-se na primeira (*M*=38.98, *DP*=9.95), com uma diferença de 11.7 pontos relativamente à pontuação média verificada na escala de *Pseudo-sintomas* (*M*=27.28, *DP*=15.22).

Tabela 8. Análise descritiva (médias, desvios-padrão e amplitude) dos resultados no SRSI na condição "simulação"

SRSI	M	DP	Min.	Máx.
Subescalas Sintomas Genuínos				
Sintomatologia Cognitiva	7.58	2.55	0	10
Sintomatologia Depressiva	7.24	2.16	2	10
Sintomatologia de Dor	7.75	3.07	0	10
Sintomatologia Somática	8.68	2.11	0	10
Sintomatologia Ansiosa	7.80	2.04	2	10
Total Sintomas Genuínos	38.98	9.95	5	50
Subescalas Pseudo-Sintomas				
Pseudo-Sintomas Cognitivos	5.65	3.25	0	10
Pseudo-Sintomas Motores	5.07	3.53	0	10
Pseudo-Sintomas Sensoriais	4.47	3.45	0	10
Pseudo-Sintomas de Dor	5.84	3.63	0	10
Pseudo-Sintomas de Ansiedade/Depressão	6.25	3.06	0	10
Total Pseudo-Sintomas	27.28	15.22	0	50

a. Comparação entre as médias alcançadas no presente estudo (escalas do SRSI) e as obtidas nos estudos de Pinheiro (2019) e Huhnt (2013)

Foi aplicado o teste *t de Student* para a verificação e análise da existência de diferenças estatisticamente significativas nos resultados médios obtidos na amostra prisional em estudo (condição "honesta") e nas amostras de referência: o estudo realizado por Pinheiro (2019), no Estabelecimento Prisional de Coimbra (cf. Tabela 9), e o estudo concretizado por Huhnt (2013), num estabelecimento prisional juvenil, tendo por base a versão alemã do SRSI (cf. Tabela 10).

Os valores inerentes à comparação entre o presente e o estudo de Pinheiro (2019), que podem ser consultados na Tabela 9, revelam que, apesar do total de ambas as escalas da amostra do presente estudo ser mais elevado do que o verificado em Pinheiro (2019), não se verificaram diferenças estatisticamente significativas. Tal é válido tanto para a escala de *Sintomas Genuínos* [$t(99)=.714$; $p=.477$], como para a escala de *Pseudo-sintomas* [$t(99)=.574$; $p=.568$].

Tabela 9. Comparação das médias obtidas no presente estudo e no de Pinheiro (2019)

SRSI	Amostra Prisional EPG M (DP) *n*= 100	Amostra Prisional Pinheiro (2019) M (DP) *n*= 85	*t*(gl= 99)
Escala Sintomas Genuínos	15.30 (10.92)	14.52 (8.91)	.714
Escala Pseudo-sintomas	3.62 (6.10)	3.27 (4.32)	.574

A análise dos resultados médios obtidos na presente amostra prisional e na amostra do estudo de Huhnt (2013) revelam diferenças estatisticamente significativas. Conforme observado na Tabela 10, os valores obtidos confirmam a existência de diferenças significativas entre as amostras, tanto para a escala de *Sintomas Genuínos* [$t(99)=-4.486$; $p<.05$], como para a escala de *Pseudo-sintomas* [$t(99)=-5.704$; $p<.05$].

Tabela 10. Comparação de médias obtidas no presente estudo e no estudo de Huhnt (2013)

SRSI	Amostra Prisional EPG M (DP) *n*= 100	Amostra Prisional Huhnt (2013) M (DP) *n*= 65	*t*(gl= 99)
Escala Sintomas Genuínos	15.30 (10.92)	20.2 (9.5)	-4.486*
Escala Pseudo-sintomas	3.62 (6.10)	7.1 (7.8)	-5.704*

* $p < .05$

i. Hipótese 1: São esperados valores mais elevados no SRSI, na condição de cenário de "simulação", comparativamente à condição de resposta "honesta"

Foi aplicado o teste *t de Student* para verificar a presença de diferenças estatisticamente significativas nos resultados médios obtidos na condição "honesta" e na condição "simulação", do SRSI (cf. Tabela 11). Os valores obtidos confirmam a existência de diferenças significativas entre ambas as condições: os valores na condição "simulação", comparativamente à condição "honesta", são todos significativamente mais elevados ($p<.001$), evidenciando um tamanho de efeito correspondente grande (d de Cohen >.80, em todas as subescalas do instrumento). Confirma-se, desta forma, a hipótese enunciada.

Tabela 11. Comparação entre os valores obtidos na condição "honesta" e na condição "simulação" (SRSI)

SRSI	Condição "honesta" M	Condição "honesta" DP	Condição "simulação" M	Condição "simulação" DP	*t*(gl=99)	*d* de Cohen
Subescalas Sintomas Genuínos						
Sintomatologia Cognitiva	2.46	2.79	7.58	2.55	-15.045***	-1.50
Sintomatologia Depressiva	3.23	2.06	7.24	2.16	-15.113***	-1.51
Sintomatologia de Dor	2.36	2.78	7.75	3.07	-14.610***	-1.47
Sintomatologia Somática	2.79	2.77	8.68	2.11	-17.656***	-1.77
Sintomatologia Ansiosa	4.48	2.50	7.80	2.04	-12.080***	-1.21
Total Sintomas Genuínos	15.37	10.95	38.98	9.95	-17.367***	-1.75
Subescalas Pseudo-Sintomas						
Pseudo-Sintomas Cognitivos	0.79	1.66	5.65	3.25	-14.053***	-1.41
Pseudo-Sintomas Motores	0.56	1.06	5.07	3.53	-12.718***	-1.27
Pseudo-Sintomas Sensoriais	0.63	1.30	4.47	3.45	-10.888***	-1.09
Pseudo-Sintomas de Dor	0.65	1.34	5.84	3.63	-14.528***	-1.45
Pseudo-Sintomas de Ansiedade/Depressão	0.99	1.67	6.25	3.06	-15.492***	-1.55
Total Pseudo-Sintomas	3.62	6.10	27.28	15.22	-15.138***	-1.51

*** $p < .001$

3. Exame da validade das pontuações do SRSI, através da análise das relações existentes entre aquelas e as pontuações de outros instrumentos psicológicos (SIMS, EVS-2 e BSI)

Num primeiro momento, serão apresentados os resultados obtidos na correlação entre a escala *Pseudo-sintomas*, do SRSI, e os restantes instrumentos de avaliação que se propõem medir constructos semelhantes àquela (comportamentos de simulação), designadamente o SIMS e a EVS-2. Neste sentido, para averiguar em que medida as pontuações do SRSI e dos restantes instrumentos estão relacionadas, foram analisadas as correlações que se estabeleceram através do coeficiente de *Pearson*.

De acordo com Mâroco (2014), a intensidade da correlação pode ser quantificável, podendo considerar-se fraca (quando o valor de *r* é inferior a .25), moderada (quando o valor de *r* recai entre .25 e .50), forte (quando o valor está compreendido entre .50 e .75) ou, por fim, muito forte (quando os valores são superiores a .75).

ii. Hipótese 2: No domínio da validade convergente, são esperadas correlações positivas elevadas entre os resultados na escala de *Pseudo-sintomas,* do SRSI, e no SIMS e na EVS-2

A relação entre a escala *Pseudo-sintomas*, do SRSI, e o SIMS pode ser observada a partir da análise da Tabela 12. No que respeita à relação entre as subescalas da escala *Pseudo-sintomas* e o total do SIMS, é de referir que todas se correlacionam positivamente, observando-se valores muito fortes (.795 e

.767) e fortes (.605, .705 e .713). De acordo com o esperado, uma vez que propõem medir constructos semelhantes, o valor de correlação mais elevado verifica-se entre o resultado total da escala *Pseudo-sintomas* e o Total do SIMS, apresentando uma associação positiva muito forte (*r*=.829).

Tabela 12. Coeficientes de *Pearson* entre a escala *Pseudo-sintomas* (SRSI) e o SIMS

SRSI	SIMS					
	Total	**AF**	**P**	**LI**	**AM**	**NI**
Subescalas Pseudo-Sintomas						
Pseudo-Sintomas Cognitivos	.795**	.652***	.395**	.413**	.772***	.759***
Pseudo-Sintomas Motores	.767**	.634***	.463**	.409**	.705***	.683***
Pseudo-Sintomas Sensoriais	.605**	.428**	.321*	.207	.652***	.706***
Pseudo-Sintomas de Dor	.713**	.613***	.342*	.421**	.706***	.572***
Pseudo-Sintomas de Ansiedade/Depressão	.705**	.692***	.393**	.324*	.625***	.522***
Total Pseudo-Sintomas	.829**	.707***	.440**	.410**	.794***	.739***

* $p < .05$ ** $p < .01$ *** $p < .001$

Legenda: AF= Perturbações Afetivas; P= Psicose; LI= Capacidade Intelectual Reduzida; AM= Perturbações Mnésicas; NI= Défice Neurológico.

Os resultados obtidos indicam correlações significativamente fortes e muito fortes entre as subescalas de *Pseudo-sintomas* e as dimensões da EVS-2 (com exceção das relações entre a dimensão *Problemas Cognitivos*, da EVS-2, e as subescalas *Pseudo-sintomas Cognitivos* e *Pseudo-sintomas Sensoriais*, que se afiguram moderadas), tal como se verifica para com as pontuações totais da EVS-2 (cf. Tabela 13).

Foram igualmente calculadas as correlações entre o resultado total de *Pseudo-sintomas*, do SRSI, e o resultado total da EVS-2, obtendo-se, tal como esperado, uma correlação significativamente muito forte: .854 (p<.01).

Tabela 13. Coeficientes de *Pearson* entre a escala *Pseudo-sintomas* (SRSI) e a EVS-2

	EVS-2			
	PE	**P**	**PC**	**Total**
Escala de Pseudo-sintomas				
Pseudo-Sintomas Cognitivos	.722***	.773***	.450**	.773**
Pseudo-Sintomas Motores	.531***	.654***	.533***	.664**
Pseudo-Sintomas Sensoriais	.670***	.777***	.448**	.753**
Pseudo-Sintomas de Dor	.634***	.657***	.604***	.727**
Pseudo-Sintomas de Ansiedade/ Depressão	.775***	.634***	.549***	.761**
Total Pseudo-Sintomas	.781***	.808***	.593***	.854**

* $p < .05$ ** $p < .01$ *** $p < .001$

Legenda: PE= Perturbações Emocionais; P= Psicose; PC= Perturbações Cognitivas

Mediante os resultados apresentados, confirma-se a Hipótese 2, na medida em que a escala *Pseudo-sintomas*, do SRSI, evidenciou correlações elevadas com os resultados totais do SIMS e da EVS-2, instrumentos (TVS, especificamente) que se destinam à avaliação do mesmo constucto teórico. Desta forma, é possível afirmar que a escala *Pseudo*-sintomas é detentora de

validade convergente.

iii. Hipótese 3: Ainda no âmbito da validade convergente, são expectáveis relações positivas elevadas entre a escala de *Sintomas Genuínos* do SRSI e resultados globais no BSI

A Tabela 14 é indicativa das correlações obtidas entre a escala *Sintomas Genuínos*, do SRSI, e os resultados do BSI.

Num primeiro momento, é de referir que os resultados obtidos pelos participantes da presente investigação nas dimensões do BSI apontam para correlações moderadas (.409 e .477, relativas às dimensões *Ideação Paranóide* e *Ansiedade Fóbica*, respetivamente) e fortes (valores que variam entre .504 e .740), com a escala *Sintomas Genuínos* do SRSI. Por sua vez, os diferentes Índices do BSI evidenciam diferentes graus de correlação com as subescalas de *Sintomas Genuínos*, tal como com a escala total. De notar que os índices *IGS* e *TSP* apresentam correlações fortes, tanto com as subescalas, como com a escala total de *Sintomas Genuínos*, com exceção do verificado na correlação moderada entre o *índice TSP* e a subescala *Sintomatologia de Dor* do SRSI (*r*=.466). Por sua vez, o *ISP* evidencia um conjunto de correlações moderadas relativamente às subescalas de *Sintomas Genuínos* e à escala total.

Deste modo, a presente hipótese apenas se revela corroborada para os índices *IGS* e *TSP*, cujas correlações com o valor total da escala de *Sintomas Genuínos* são .747 e .662, respetivamente.

Tabela 14. Coeficientes de *Pearson* entre a escala *Sintomas Genuínos* (SRSI) e os índices globais do BSI

	BSI		
	IGS	**TSP**	**ISP**
Escala Sintomas Genuínos			
Sintomatologia Cognitiva	.616**	.557**	.388**
Sintomatologia Depressiva	.656**	.576**	.468**
Sintomatologia de Dor	.549**	.466**	.359**
Sintomatologia Somática	.700**	.626**	.419**
Sintomatologia Ansiosa	.653**	.588**	.449**
Total Sintomas Genuínos	.747**	.662**	.419**

* $p < .05$ ** $p < .01$ *** $p < .001$

Legenda: IGS= Índice Geral de Sintomas; TSP= Total de Sintomas Positivos; ISP= Índice de Sintomas Positivos.

iv: Hipótese 4: No contexto da validade discriminante, são esperadas relações positivas mais reduzidas entre os resultados das escalas/instrumentos que avaliam constructos distintos

Ao contrário do desejável e esperado no âmbito da Hipoótese 4, a correlação entre o resultado total de *Sintomas Genuínos* e o resultado total do SIMS foi estatisticamente significativa e forte (*r*=.749; cf. Tabela 15). Por sua vez, é possível constatar que o total da escala de *Sintomas Genuínos* do SRSI

apresenta correlações muito fortes e fortes com as subsescalas do SIMS, à exceção das subescalas P e LI, cuja correlação registada foi moderada (.403 e .294, respetivamente). No entanto, verificaram-se correlações fracas entre a subescala LI, do SIMS, e as subescalas *Sintomatologia Cognitiva* (*r*=.144), *Sintomatologia Somática* (*r*=.242) e *Sintomatologia Ansiosa* (*r* =.122), do SRSI, assim como entre a subescala P, pertencente ao SIMS, e as subescalas *Sintomatologia Cognitiva* (*r*=.235) e *Sintomatologia Ansiosa* (*r*=.213), do SRSI.

Tabela 15. Coeficientes de *Pearson* entre a escala *Sintomas Genuínos* (SRSI) e as escalas do SIMS

SRSI	SIMS					
	Total	**AF**	**P**	**LI**	**AM**	**NI**
Subescalas Sintomas Genuínos						
Sintomatologia Cognitiva	.554**	.459**	.235**	.144**	.654**	.534**
Sintomatologia Depressiva	.720**	.701**	.356**	.327**	.663**	.559**
Sintomatologia de Dor	.690**	.489**	.504**	.383**	.700**	.590**
Sintomatologia Somática	.679**	.606**	.340**	.242**	.729**	.558**
Sintomatologia Ansiosa	.523**	.551**	.213**	.122**	.469**	.496**
Total Sintomas Genuínos	.749**	.676**	.403**	.294**	.787**	.668**

* $p < .05$ ** $p < .01$ *** $p < .001$

Legenda: AF= Perturbações Afetivas; P= Psicose; LI= Capacidade Intelectual Reduzida; AM= Perturbações Mnésicas; NI= Défice Neurológico.

Um panorama idêntico de resultados inesperados verifica-se nas relações entre o total da escala *Sintomas Genuínos* (e inerentes subescalas), do SRSI, e as dimensões da EVS-2, verificando-se correlações significativas, fortes e muito fortes (cf. Tabela 16). Não obstante, a dimensão PC, da EVS-2, relacionou-se moderadamente com as subescalas *Sintomatologia Cognitiva* (.445), *Sintomatologia Depressiva* (.463), *Sintomatologia de Dor* (.468) e *Sintomatologia Ansiosa* (.487). Assim, rejeita-se o enunciado pela hipótese.

Tabela 16. Coeficientes de *Pearson* entre a escala *Sintomas Genuínos* (SRSI) e as dimensões da EVS-2

	EVS-2			
	PE	**P**	**PC**	**Total**
Escala Sintomas Genuínos				
Sintomatologia Cognitiva	.704**	.602**	.445**	.688**
Sintomatologia Depressiva	.808**	.656**	.463**	.760**
Sintomatologia de Dor	.621**	.731**	.468**	.718**
Sintomatologia Somática	.733**	.761**	.517**	.792**
Sintomatologia Ansiosa	.750**	.627**	.487**	.730**
Total Sintomas Genuínos	.822**	.776**	.546**	.843**

* $p < .05$ ** $p < .01$ *** $p < .001$

Legenda: PE= Perturbações Emocionais; P= Psicose; PC= Perturbações Cognitivas

Ao analisar a Tabela 17, é possível observar as correlações que, de forma não expectável, se verificaram entre os índices globais do BSI (assim

como nas suas dimensões – não se encontram reportadas na Tabela 17) e as subescalas e escala total de *Pseudo-sintomas*: valores a oscilarem entre moderados e fortes (à exceção da relação verificada entre a dimensão *Ideação Paranóide*, do BSI, e subescala *Pseudo-sintomas Motores*, do SRSI, que se revelou fraca – *r*=.198).

Tabela 17. Coeficientes de *Pearson* entre a subescala *Pseudo-sintomas* (SRSI) e os índices globais do BSI

	BSI		
	IGS	**TSP**	**ISP**
Escala de Pseudo-sintomas			
Pseudo-Sintomas Cognitivos	.561**	.450**	.361**
Pseudo-Sintomas Motores	.492**	.407**	.307**
Pseudo-Sintomas Sensoriais	.642**	.512**	.400**
Pseudo-Sintomas de Dor	.513**	.383**	.389**
Pseudo-Sintomas de Ansiedade/ Depressão	.601**	.519**	.356**
Total Pseudo-Sintomas	.652**	.528**	.388**

* $p < .05$ ** $p < .01$ *** $p < .001$
Legenda: IGS= Índice Geral de Sintomas; TSP= Total de Sintomas Positivos; ISP= Índice de Sintomas Positivos.

Ainda no âmbito da Hipótese 4, analisámos as correlações existentes entre as pontuações de cada condição ("honesta" e de "simulação"), relativamente a cada escala do SRSI – cf. Tabela 18; e entre as pontuações da escala de *Sintomas Genuínos*, da condição "honesta", e da escala de *Pseudo-sintomas*, da condição "simulação" – cf. Tabela 19. Na globalidade, os resultados foram ao encontro do esperado: correlações positivas, mas baixas.

Como mencionado, a Tabela 18 evidencia, na sua globalidade, correlações positivas baixas (com valores a variar entre .003 e .299). Porém, apresenta algumas exceções ao referido previamente, entre elas: algumas correlações negativas em ambas as escalas (e.g., correlação entre a subescala *Sintomatologia Ansiosa* [condição "simulação"] e as subescalas *Sintomatologia Cognitiva* e *Sintomatologia Somática* [condição "honesta"]; correlação entre a subescala *Sintomatologia de Dor* [condição "simulação"] e a subescala *Sintomatologia Depressiva* [condição "honesta"]; correlação entre a o resultado total da escala *Pseudo-sintomas* [condição "simulação"] e a subescala *Pseudo-sintomas Cognitivos* [condição "honesta"]); e algumas correlações significativas ($p < .05$ e $p < .01$; e.g., correlação entre a subescala *Sintomatologia Depressiva* [condição "simulação"] e as subesacalas *Sintomatologia Depressiva* e *Sintomatologia Ansiosa* [condição "honesta"]).

Tabela 18. Coeficiente de *Pearson* entre os resultados das escalas do SRSI, na condição "honesta" *versus* na condição "simulação"

SRSI (Condição "honesta")	SRSI (Condição "simulação")					
Sintomat. Genuína	**Cognitiva**	**Depressiva**	**De Dor**	**Somática**	**Ansiosa**	**Total**
Cognitiva	.189	.044	.115	.102	-.035	.180
Depressiva	.035	.208*	-.086	.107	.029	.137
De Dor	.167	.066	.216*	.126	.003	.170
Somática	.106	.033	.086	.084	-.049	.158
Ansiosa	.206*	.252*	.148	.299*	.282*	.047
Total	.104	.104	.142	.065	.260**	.165
Pseudo-Sintomas	**Cognitivos**	**Motores**	**Sensoriais**	**Dor**	**Ansiedade / Depressão**	**Total**
Cognitivos	.128	.006	.106	.129	.038	-.065
Motores	.107	.171	.124	.179	.131	.152
Sensoriais	.108	.129	.129	.178	.107	.146
Dor	.092	.131	.132	.225*	.149	.163
Ansiedade/ Depressão	.063	.017	.062	.048	.060	.055
Total	.114	.087	..124	.167	.105	.133

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

É de referir a existência de correlações positivas fracas ou nulas, mas não significativas, entre as subescalas e o total da escala de *Sintomas Genuínos* (condição "honesta") e as subescalas e o total de *Pseudo-sintomas* (condição "simulação"), variando os valores entre .005 e .135. Todavia, a subescala *Sintomatologia Ansiosa*, da condição "honesta", revelou correlacionar-se de forma negativa com todas as subescalas de *Pseudo-sintomas* (cf. Tabela 19).

Tabela 19. Coeficiente de *Pearson* entre os resultados da escala *Sintomas Genuínos* (condição "honesta") e de *Pseudo-sintomas* (condição "simulação")

Sintomas Genuínos (Condição "Honesta")	Pseudo-Sintomas (Condição "Simulação")					
	Cognitivos	Motores	Sensoriais	Dor	Ansiedade/ Depressão	Total
Sintomatologia Cognitiva	.019	.058	.037	.099	.047	.059
Sintomatologia Depressiva	.053	.067	.059	.005	.055	.052
Sintomatologia de Dor	.111	.049	.109	.135	.046	.101
Sintomatologia Somática	.049	.039	.081	.067	.048	.063
Sintomatologia Ansiosa	-.097	-.016	-.002	-.027	-.028	-.037
Total	.033	.046	.068	.071	.040	.058

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

iv: Hipótese 5: No âmbito da validade discriminante, são esperadas diferenças estatisticamente significativas entre as pontuações obtidas em ambas as condições do SRSI

Ao examinar a presente Hipótese, averiguou-se a presença de diferenças significativas entre as condições "honesta" e de "simulação". Para o efeito, recorremos ao teste *t de Student*, que permitiu observar valores estatisticamente mais elevados (p<.001) na condição "simulação", comparativamente aos verificados na condição "honesta" (cf. Tabela 20), tanto para a escala *Sintomas Genuínos*, como para a escala *Pseudo-sintomas*, sendo o tamanho de efeito correspondente grande (*d* de Cohen >.80).

Tabela 20. Comparação entre as pontuações obtidas na condição "honesta" e na condição "simulação" (SRSI)

SRSI	Condição "honesta" M (DP)	Condição "honesta" Ampl.	Condição "simulação" M (DP)	Condição "simulação" Ampl.	*t*(gl=99)	*d* de Cohen
Escala Sintomas Genuínos						
Sintomatologia Cognitiva	2.46 (2.79)	0 – 10	7.58 (2.55)	0 – 10	-15.045***	-1.50
Sintomatologia Depressiva	3.23 (2.06)	0 – 9	7.24 (2.16)	2 – 10	-15.113***	-1.51
Sintomatologia de Dor	2.36 (2.78)	0 – 10	7.75 (3.07)	0 – 10	-14.610***	-1.47
Sintomatologia Somática	2.79 (2.77)	0 – 10	8.68 (2.11)	0 – 10	-17.656***	-1.77
Sintomatologia Ansiosa	4.48 (2.50)	0 – 10	7.80 (2.04)	2 – 10	-12.080***	-1.21
Total Sintomas Genuínos	15.37 (10.95)	1 – 42	38.98 (9.95)	0 – 50	-17.367***	-1.75
Escala Pseudo-Sintomas						
Pseudo-Sintomas Cognitivos	0.79 (1.66)	0 – 7	5.65 (3.25)	0 – 10	-14.053***	-1.41
Pseudo-Sintomas Motores	0.56 (1.06)	0 – 4	5.07 (3.53)	0 – 10	-12.718***	-1.27
Pseudo-Sintomas Sensoriais	0.63 (1.30)	0 – 9	4.47 (3.45)	0 – 10	-10.888***	-1.09
Pseudo-Sintomas de Dor	0.65 (1.34)	0 – 7	5.84 (3.63)	0 – 10	-14.528***	-1.45
Pseudo-Sintomas de Ansiedade/Depressão	0.99 (1.67)	0 – 9	6.25 (3.06)	0 – 10	-15.492***	-1.55
Total Pseudo-Sintomas	3.62 (6.10)	0 – 29	27.28 (15.22)	0 – 50	-15.138***	-1.51

*** p < .001

iv: Hipótese 6: No contexto da validade discriminante, são esperadas diferenças estatisticamente significativas entre as pontuações obtidas pelos participantes reclusos *versus* não-reclusos (amostra de controlo)

Ainda no que concerne ao exame da validade discriminante, averiguou-se a presença de diferenças estatisticamente significativas entre as pontuações obtidas pelos participantes da presente amostra prisional (reclusos) e aquelas que foram obtidas pela amostra de controlo (não-reclusos), constituída pelos participantes dos estudos de Dwarkadas (2018) e Domingues (2019). Para tal, recorremos ao teste *t de Student*, que permitiu observar valores estatisticamente mais elevados (p<.05) na presente amostra prisional, quando comparados com os obtidos pela amostra de controlo (cf. Tabela 21), sendo o

tamanho de efeito correspondente grande (*d* de Cohen >.80). Tal verificou-se nas escalas globais do SRSI, assim como nas suas subescalas (à exceção das subescalas *Sintomatologia Cognitiva*, *Sintomatologia Depressiva* e *Sintomatologia Somática*, cujas diferenças não se revelaram significativas).

Tabela 21. Comparação entre as pontuações obtidas (no SRSI) pela presente amostra prisional e na amostra de controlo

SRSI	Amostra prisional *n*=100	Amostra de controlo *n*=156	*t*	*d* de Cohen
	M (DP)	M (DP)		
Escala Sintomas Genuínos				
Sintomatologia Cognitiva	2.46 (2.79)	2.07 (2.59)	1.158	1.16
Sintomatologia Depressiva	3.23 (2.06)	1.90 (1.86)	5.356	5.35
Sintomatologia de Dor	2.36 (2.78)	1.04 (1.88)	4.129***	4.13
Sintomatologia Somática	2.79 (2.77)	2.84 (2.50)	-.149	-0.15
Sintomatologia Ansiosa	4.48 (2.50)	2.13 (2.14)	7.754**	7.76
Total Sintomas Genuínos	15.37 (10.95)	9.97 (8.53)	4.134 **	4.14
Escala Pseudo-Sintomas				
Pseudo-Sintomas Cognitivos	0.79 (1.66)	0.53 (1.22)	1.374*	1.38
Pseudo-Sintomas Motores	0.56 (1.06)	0.21 (0.70)	2.912***	2.90
Pseudo-Sintomas Sensoriais	0.63 (1.30)	0.43 (0.81)	1.379*	1.39
Pseudo-Sintomas de Dor	0.65 (1.34)	0.24 (0.67)	2.851***	2.85
Pseudo-Sintomas de Ansiedade/Depressão	0.99 (1.67)	0.35 (0.85)	3.578***	3.58
Total Pseudo-Sintomas	3.62 (6.10)	1.75 (3.28)	2.815***	2.82

* $p < .05$ ** $p < .01$ *** $p < .001$

4. Frequência das respostas obtidas às questões integrantes do Questionário Pós-experimental de verificação de manipulação

A Tabela 22 informa acerca da forma e frequência com que os participantes do presente estudo responderam ao Questionário Pós-experimental elaborado. Para a primeira questão – "*Considera este cenário compreensível?*" –, a percentagem de resposta mais elevada remete para o item "*Sim*" (*n*=98; 98%). Já na questão "*Considera este cenário viável e realista?*", a maioria dos reclusos-participantes respondeu "*Sim*" (*n*=65; 65%), verificando-se, inversamente, uma frequência mais elevada na resposta "*Não*" (*n*=53; 53%), no que respeita ao item "*Achou compreensível que o personagem simulasse sintomas?*". Relativamente às duas últimas questões ("*Foi fácil para si imaginar que era o personagem principal?*" e "*Fez o seu melhor para comunicar os sintomas de modo convincente e credível?*"), a maioria dos participantes respondeu "*Sim*" (*n*=66; 66% e *n*=90; 90%, respetivamente). A tendência de respostas inerente à última questão revela que uma maioria de participantes depositou esforço no exercício de responder à situação requerida na *cover story*.

Tabela 22. Frequência das respostas às questões do Questionário Pós-experimental

	N	%
1) "*Considera este cenário compreensível* (*ao nível da compreensão do que está a ser pedido)*?"		
Sim	98	98.0
Não	0	0
Moderadamente	2	2.0
2) "*Considera este cenário viável e realista*?"		
Sim	65	65.0
Não	26	26.0
Moderadamente	9	9.0
3) "*Achou compreensível que o personagem simulasse sintomas*?"		
Sim	33	33
Não	53	53.0
Moderadamente	14	14.0
4) "*Foi fácil para si imaginar que era o personagem principal?*"		
Sim	66	66.0
Não	18	18.0
Moderadamente	16	16.0
5) "*Fez o seu melhor para comunicar os sintomas de modo convincente e credível*?"		
Sim	90	90.0
Não	1	1.0
Moderadamente	9	9.0

V - Discussão dos resultados

O exame da credibilidade e validade do autorrelato, sobretudo das condutas da simulação e do exagero de sintomatologia, é uma componente incontornável e de extrema relevância a considerar na avaliação (neuro)psicológica, tanto em contexto forense, como clínico (Kemp et al., 2008). Para concretizar esse objetivo, é necessária a existência de vários instrumentos que potenciem a deteção de atitudes de exagero de sintomas (TVS), como por exemplo o SIMS, a EVS-2 e, também, o SRSI.

O presente estudo teve como principal objetivo continuar o processo de validação do SRSI para o contexto prisional português – processo previamente iniciado por Pinheiro (2019) –, instrumento que poderá constituir-se como uma estratégia adicional de avaliação da credibilidade dos sintomas reportados nos contextos de avaliação forense (e, no caso particular deste estudo, criminal/ prisional). Para tal, e com base numa amostra de 100 reclusos-participantes, recorremos a um *design* de simulação análogo, que permitiu comparar respostas fornecidas ao SRSI numa condição de "resposta honesta" e numa condição de "resposta de simulação".

Inicialmente, são de valorizar os valores de consistência interna obtidos para as escalas e subescalas do SRSI para a condição "honesta", que se

refletiram em valores elevados (.94 e .93 para a escala *Sintomas Genuínos* e de *Pseudo-sintomas*, respetivamente). Os referidos valores assemelham-se àqueles que se apuraram no estudo com uma amostra prisional levado a cabo pelo alemão Huhnt (2013), cujos valores de *alfa* foram de .90, para a escala *Sintomas Genuínos*, e de .92, para a escala de *Pseudo-sintomas*. Já nos estudos realizados em Portugal, sobretudo naquele que se concretizou em contexto prisional, Pinheiro (2019) obteve valores de *alfa* muito bons, tanto para a escala *Sintomas Genuínos* (.80), como para a escala de *Pseudo-sintomas* (.80).

A condição "simulação" refletiu valores de consistência interna igualmente muito elevados para ambas as escalas do instrumento (escala de *Sintomas Genuínos* com um valor de .94 e a escala de P*seudo-sintomas* com .98). Uma vez comparando os resultados descritos com aqueles que foram obtidos na investigação conduzida por Domingues (2019), cujos valores de *alfa de Cronbach* foram respeitáveis (.72 na escala *Sintomas Genuínos*) e muito bons (.88 na escala de *Pseudo-sintomas*) para a condição "simulação", é relevante ressaltar a noção de que, geralmente, valores de *alfa* superiores a .90 podem refletir uma elevada e/ou exagerada homogeneidade dos itens do instrumento. Desta forma, e alargando a reflexão aos valores de *alfa* igualmente elevados para a condição "honesta", é de assinalar que diversas são as variáveis que podem justificar um valor de *alfa de Cronbach* extremamente elevado, sem que isso reflita, necessariamente, uma consistência interna elevada. Entre elas, identificam-se: questionários muito longos ou amostras muito numerosas (Bland & Altman, 1997); a adoção de um enviesamento de resposta conhecido como *efeito de aquiescência* (tendência de responder afirmativamente aos itens, desconsiderando o seu conteúdo descritivo; Billiet & McClendon, 2000); ou um conteúdo de itens redundante (o que poderia ser apurado através da análise das correlações inter-item, que seriam muito elevadas; Cattel, 1972). No seguimento deste último, é de valorizar a sensação (bastante comentada pelos participantes ["*Mas isto só fala de dores*?" ou "*Tanta repetição*!" – sic]) de que os itens do instrumento recaem, constantemente, sobre o(s) mesmo(s) constructo(s).

Por sua vez, foram apurados valores de *alfa de Cronbach* problemáticos para algumas escalas que integram o SIMS: a escala LI, NI e P, cujos valores se revelaram inaceitáveis. Num processo de comparação entre outros estudos investigativos, apesar de, no geral, se considerar que a consistência interna do resultado total do SIMS é satisfatória (e.g., α=.80 em Cima et al., 2003; α=.81 em Mota, 2008; α=.825 no atual estudo), alguns apresentam valores de *alfa* menos desejáveis para as escalas do instrumento. É o caso dos estudos de Mota (2008) e de Merckelbach e Smith (2003), nos quais se observaram valores considerados inaceitáveis, designadamente para as escalas LI e NI. Embora tal reflita que algumas escalas podem revelar-se problemáticas, é de salientar que as mesmas não foram delineadas para, *per si* e de forma autónoma, detetarem a simulação de psicopatologia (apesar de permitirem uma análise qualitativa importante, nomeadamente quando a pontuação total excede o ponto de corte selecionado; Impelen et al., 2014).

No âmbito da análise das estatísticas descritivas inerentes às pontuações obtidas no SRSI, em ambas as condições, foi possível analisar um conjunto

de dados que permitiram testar a Hipótese 1, corroborando-a. O racional que subjaz ao instrumento fez-nos prever que a escala de *Sintomas Genuínos* seria aquela com mais sintomas reportados, sobretudo na condição "honestidade", uma vez que se pressupunha estarmos perante respondentes sem motivações externas para adotar um estilo de resposta enviesado (ou seja, sem motivações para apresentar um número elevado de pseudo-sintomas). Tal reflete os resultados dos estudos internacionais (Boskovic et al., 2019; Helvoort et al., 2019; Merten et al., 2016) e dos estudos portugueses de Dwarkadas (2018), Pinheiro (2019) e Domingues (2019). Também na presente investigação, a escala de *Sintomas Genuínos* revelou ser a escala com um maior número de sintomas reportados (quer na condição "honestidade", quer na condição "simulação"), revelando-se o aumento estatisticamente significativo (p <.001; dados igualmente observados no estudo de Domingues, 2019; Boskovic et al., 2019; Helvoort et al., 2019), ao qual se associa uma tamanho de efeito grande (d de Cohen >.80). Uma possível explicação pode recair na boa compreensão, e inerente prática, do requerido na *cover story* (o que é coerente com o facto de 98 dos 100 participantes terem revelado compreender as instruções, no âmbito do QPE): os participantes foram instruídos a simular e/ou exagerar sintomas, mas alertados para a necessidade de serem cuidadosos e convincentes nessa tarefa, tendo sido explicitadas as consequências de uma prática que não atendesse a tais recomendações – a indesejável deteção da conduta de simulação. Assim, o facto de a escala *Sintomas Genuínos* integrar sintomas mais comummente observados em diferentes populações (saudáveis ou clínicas; Merten et al., 2016), e a escala de *Pseudo-sintomas* ser constituída por sintomas mais extremos e atípicos, concorre igualmente para a explicação, dado o empenho depositado, por parte dos participantes, na tarefa de exagerar e simular sintomas de forma credível. O referido é congruente com a perceção de autoeficácia expressa pelos próprios (no âmbito do QPE, 90 participantes mencionaram ter feito o seu melhor para reportar os sintomas de forma verosímil e convincente), sendo tal suportado pelos comentários que surgiram ao longo do procedimento de resposta ao SRSI, na condição "simulação" (e.g., "*Isto* [este item] *já é demasiado extremo!*" ou "*Ia ser logo descoberto se assinalasse este* [item]*!*" – sic).

Apesar de a prevalência de condutas de simulação e exagero de sintomatologia no contexto prisional não ser consensual, diversos estudos apontam para taxas de prevalência consideráveis (ainda que mais baixas do que as encontradas no contexto médico-legal, quer no âmbito civil ou penal; McDermott & Sokolov, 2009; Mittenberg et al., 2002). Neste sentido, é de evidenciar que, na condição "honesta", 18 dos 100 participantes assinalaram mais de 6 (ponto de corte de rastreio) ou 9 (ponto de corte diagnóstico) pseudo-sintomas. Nesta condição, supusemos a inexistência de motivos que justificassem a identificação, por parte do instrumento, dos 18 participantes como estando a exagerar ou simular sintomatologia. No entanto, determinados fatores podem auxiliar a sua compreensão (e.g., emocionais, motivacionais, etc; Young, 2021): alguns estudos explicam o exagero de sintomatologia por parte de reclusos como forma de colmatar o aborrecimento face ao processo de administração do protocolo de avaliação, ou meramente pela atenção e

divertimento que podem retirar dessa conduta (Denney, 2007; Wall et al., 2015). Outros estudos referem que simuladores em contexto prisional tendem a ser menos escolarizados (e.g., McDermott & Sokolov, 2009) e mais novos (Norris & May, 1998). Após a análise dos nossos dados, 11 dos participantes identificados como simuladores, pelo SRSI, concluíram 9 anos de escolaridade, 5 deles tinham 6 anos completos, 1 concluiu 12 anos e 1 concluiu 4 anos. Por sua vez, as idades variaram entre os 18 e os 62 anos. Identifica-se, ainda, um risco aumentado de simulação aquando da presença de reclusos que se encontrem em situação de prisão preventiva (ou com processos penais pendentes; Denney, 2007). Porém, não indagámos previamente os participantes acerca da sua situação jurídico-penal, pelo que não podemos inferir conclusões que suportem o referido.

Com o intuito de concretizar um dos objetivos específicos do presente estudo, foram analisados os valores da média e desvio-padrão da condição "honesta" para a amostra recolhida (*N*=100). Por sua vez, estes foram comparados com os valores obtidos no estudo português, realizado por Pinheiro (2019; *N*=85), e no estudo alemão, pragmatizado por Huhnt (2013; *N*=65). Num primeiro momento, é de sublinhar o padrão evidenciado pelas três investigações: a média da escala *Sintomas Genuínos* é consideravelmente superior ao valor da média da escala *Pseudo-sintomas*. Ademais, aquando da primeira comparação mencionada (presente estudo *versus* estudo de Pinheiro, 2019), verificou-se a inexistência de diferenças significativas entre os resultados obtidos em ambas as amostras. Contudo, o mesmo não se verificou quando comparadas as médias entre os resultados da presente amostra e os observados no estudo de Huhnt (2013), tendo este último obtido pontuações significativamente superiores, em ambas as escalas. Tal pode dever-se às características dos participantes envolvidos em ambos os estudos, nomeadamente à idade (tal como referido anteriormente, pressupõe-se que reclusos mais novos têm maior predisposição para simular sintomatologia; Norris & May, 1998): enquanto no presente estudo a média é de 40.51 anos (*DP*=12.79), o estudo de Huhnt (2013) contou com a participação de jovens reclusos, com 20.71 anos, em média (*DP*=1.73).

O presente estudo permitiu agregar evidências de que a escala *Pseudo-sintomas* possui validade convergente, uma vez que se verificaram correlações positivas muito fortes entre a mesma e os totais do SIMS e da EVS-2 (instrumentos que pretendem mensurar constructos semelhantes; AERA et al., 2014). Tal indica que os participantes cujas pontuações se revelaram elevadas na escala de *Pseudo-sintomas* obtiveram, no geral, uma pontuação elevada no SIMS e na EVS-2. No que respeita à relação com o SIMS, o presente estudo observou uma correlação muito forte e estatisticamente significativa entre o seu resultado total e a escala de *Pseudo-sintomas* (r=.829, p<.01). Todavia, e de forma não expectável, também a escala de *Sintomas Genuínos* se relacionou positiva, forte e significativamente com o resultado total do SIMS (r=.749, p<.01) – evidencia ausência de validade discriminativa. Tais resultados são congruentes com o verificado em Pinheiro (2019), que constatou valores fortes (r=.686, p<.01) e muito fortes (r=.765, p<.01) na relação entre o resultado total do SIMS e a escala *Sintomas Genuínos* e

Pseudo-sintomas, respetivamente. No mesmo sentido surgiram os resultados dos estudos de Dwarkadas (2018), Domingues (2019) e Merten e colaboradores (2016): correlações positivas, fortes e/ou muito fortes e significativas entre o resultado total do SIMS e ambas as escalas do SRSI.

No que respeita a dados provenientes de outros estudos internacionais acerca das relações observadas entre o SRSI e o SIMS, é de destacar alguns resultados distintos. Por um lado, Giger e Merten (2019) observaram, no âmbito da condição "honesta", correlações problemáticas: forte e significativa (r=.67, p<.001) entre a escala *Sintomas Genuínos* e o SIMS, mas não significativa e fraca (r=.13) no que respeita à relação entre o SIMS e a escala *Pseudo-sintomas*. Segundo os autores, tal pode dever-se ao efeito da pequena variação inerente à apresentação de pseudo-sintomas, por parte dos participantes, já que, numa condição de "simulação", as respetivas correlações aumentaram para .65 e .69, respetivamente (p<.001). Por sua vez, o estudo de Huhnt (2013) possibilitou observar correlações igualmente problemáticas: moderada (r=.49) entre o SIMS e a escala *Pseudo-sintomas*, e forte (r=.51) entre o primeiro e a escala de *Sintomas Genuínos* do SRSI. Não obstante, a correlação moderada parece ter refletido o intervalo temporal (2 a 3 semanas) entre a primeira (aplicação do SRSI) e a segunda (aplicação do SIMS) sessão.

No âmbito das relações verificadas entre o SRSI e a EVS-2, observaram-se correlações positivas, muito fortes (r=.854) e estatisticamente significativas (p<.01) entre o total da escala *Pseudo-sintomas* e o total da EVS-2 (evidenciando validade convergente), mas também entre este último e a escala de *Sintomas Genuínos* (r=.843, p <.01) – carência de validade discriminante. O referido resultado inesperado foi ao encontro do verificado no estudo de Dwarkadas (2018), no qual se observaram correlações fortes e significativas entre o resultado total da EVS-2 e ambas as escalas do SRSI. Assim, considerando as relações problemáticas e não expectáveis observadas, os resultados do presente estudo apoiam o constatado na literatura: pontuações elevadas em ambas as escalas do SRSI encontram-se associadas a resultados elevados noutros TVS (particularmente no SIMS: Domingues, 2019; Merten et al., 2016; Pinheiro, 2019; e na EVS-2: Dwarkadas, 2018).

Ainda no domínio de análise da validade convergente, a inclusão do BSI no protocolo de avaliação objetivou analisar a sua relação com a escala *Sintomas Genuínos*, do SRSI (uma vez que ambos visam identificar psicopatologia real). Na observação das correlações entre o resultado da escala *Sintomas Genuínos* e os três índices globais do BSI, verificou-se que são positivas, significativas (p <.01) e fortes, à exceção do *ISP*, cuja correlação é moderada (r= .419). Desta forma, e como expectável, valores elevados na escala *Sintomas Genuínos* estão associados à obtenção de valores elevados nos índices do BSI, sendo, assim, possível corroborar a Hipótese 3.

Não obstante, são de evidenciar as relações igualmente fortes (índices *IGS* e *TSP*), moderadas (índice *ISP*) e significativas (p <.01) verificadas entre os índices globais do BSI e a escala *Pseudo-sintomas* do SRSI (ausência de validade discriminativa), o que é indicativo de que esta última pode ser sensível à presença de psicopatologia genuína. Tal significa que, perante examinandos que apresentem psicopatologia real, estes podem ser

erroneamente identificados como alguém que recorreu à conduta de simulação de sintomatologia – Erro de Tipo I (Impelen, 2014). Estes resultados são congruentes com o verificado em estudos empíricos anteriormente realizados em contexto internacional e nacional (e.g., Boskovic et al.; 2019; Dwarkadas, 2018; Pinheiro, 2019), nos quais se observaram correlações igualmente problemáticas (moderadas e/ou fortes e significativas) entre os resultados nos instrumentos de avaliação de psicopatologia utilizados (e.g., BSI-18; BDI-II, STAI-Y e BPI-SF) e a escala de *Pseudo-sintomas* do SRSI. Desta forma, o referido conjunto de resultados vem contrapor o salientado por Helvoort e colaboradores (2019), que concluíram que a escala de *Pseudo-sintomas* havia sido insensível à psicopatologia de pacientes reais e/ou genuínos.

Para além da sensibilidade à psicopatologia genuína verificada na escala de *Pseudo-sintomas*, outras variáveis que podem concorrer para a ocorrência do erro de Tipo I (Young, 2021), a saber: capacidades cognitivas insuficientes, fadiga ou uma compreensão diminuída – variável a considerar particularmente em estudos com população prisional, uma vez que a maioria dos participantes não é elevadamente escolarizado. É neste contexto que se afigura relevante identificar alguns itens que aparentaram alguma controvérsia em contexto prisional, já que a sua apresentação pode identicamente potenciar uma sobreidentificação de sintomas, sem que a resposta do participante evidencie, necessariamente, a sua presença verídica. Exemplifique-se: no item 11 ("*De manhã, acordo mais cedo do que o normal*"), os participantes tenderam a comparar a sua atual situação de reclusão, com a rotina em liberdade; nos itens 93 e 98 ("*Cometi muitos erros ao longo da minha vida*"; "*Sou o culpado da minha própria infelicidade*"), suspeita-se que podem ter sido afetados pela desejabilidade social, na medida em que pessoas reclusas tendem a assumir (social e/ou legalmente) os factos pelos quais estão condenados, pelas vantagens percecionadas nessa ação. Nestas circunstâncias, uma resposta positiva (verificada na maioria dos participantes para os três itens) pode não ser indicadora de sintomatologia depressiva genuína. Assim, embora não se identifique ambiguidade na formulação dos itens do SRSI, e se reconheça a sua simplicidade (Merten et al., 2016), é sempre pertinente atender ao contexto no qual a avaliação se insere, de forma a analisar o impacto que alguns itens acarretam no diagnóstico de simulação, sem que tal evidencie, necessariamente, simulação de psicopatologia. Também Pinheiro (2019) identificou os itens enunciados como potencialmente problemáticos para a população prisional, evidenciando, adicionalmente, a controvérsia inerente ao item 59 ("*Quando saio de casa, imagino tudo o que de terrível pode acontecer*"), já que os participantes se encontram em situação de reclusão. Neste sentido, apesar de, na globalidade, os participantes do presente estudo não terem contestado o conteúdo do item 59, seria pertinente reformulá-lo e adequá-lo a futuras investigações em meio prisional.

Prosseguindo no âmbito da validade discriminante (Hipótese 4), previa-se igualmente a existência de relações positivas mais reduzidas entre as escalas globais do SRSI. Tal fora constatado no presente estudo empírico, sobretudo aquando da verificação das correlações entre ambas as condições, no âmbito da escala de *Sintomas Genuínos*, por um lado, e da escala de *Pseudo-sintomas*,

por outro. Também a relação entre a escala *Sintomas Genuínos* (condição "honesta") e a escala *Pseudo-sintomas* (condição "simulação") foi ao encontro do previsto. Assim, na globalidade, as referidas correlações são positivas e reduzidas, mas não significativas, o que corrobora o previsto na Hipótese 4.

Numa observação relativa à forma de validade discriminante examinada na Hipótese 4, é de referir que, segundo Campbell e Fiske (1959, como citados em Clark & Watson, 2019), aquela estabelece-se aquando da demonstração de que correlações convergentes são mais elevadas do que as discriminantes. De facto, guiando-nos pelo referido, corroboraríamos a Hipótese 4 na íntegra. Não obstante, pelo facto de o significado do termo "mais elevadas" se prever ambíguo neste contexto, é aconselhável atender a requisitos mais rigorosos, para que se possa concluir a presença de validade discriminante sem reservas (Clark & Watson, 2019). Por um lado, a correlação convergente deve ser significativamente mais alta do que os coeficientes discriminantes, relativamente aos quais é comparada (o que é mais difícil de cumprir). Por outro, e sendo o aspeto mais relevante, para deter um poder (significância) estatístico suficiente, há necessidade de amostras com, pelo menos, 200 observações, o que não se verifica na presente investigação e o que, por sua vez, reforça a rejeição de parte da Hipótese 4.

Uma vez que os contextos investigativos têm implicações distintas das que se verificam nos contextos da vida real, o facto de os resultados obtidos se reportarem à *performance* de participantes de uma investigação (e não a uma circunstância real na qual estivessem inseridos) pode ter influenciado os resultados. Desta forma, por melhor descritos que os cenários se encontrem, a conduta dos participantes pode não refletir exatamente a que seria adotada por um indivíduo inserido veridicamente na situação descrita (fraca validade externa – limitação do DSA; Rogers & Cruise, 1998).

Não obstante o referido, é relevante reforçar o padrão que se tem vindo a verificar no âmbito desta tipologia de investigação análoga (DAS), intra ou inter-sujeitos (cf. Domingues, 2019; Helvoort et al., 2019; Boskovic et al., 2019): na condição "simulação", comparativamente à condição "honestidade", os indivíduos apresentam significativamente mais sintomas em ambas as escalas do SRSI. No presente estudo, especificamente, na condição "honestidade", os participantes apresentaram, em média, poucos pseudo-sintomas (M=3.62; DP=6.10), ficando esse valor abaixo do ponto de corte >6 (de rastreio) e >9 (de diagnóstico). Em contraste, na condição "simulação", os participantes evidenciaram um número significativamente mais elevado [t(99)=-15.138; p=<.001, d=-1.51] de pseudo-sintomas (M= 39.98; DP=27.28), tendo 88 dos 100 participantes excedido o ponto de corte de 6, e 87 o ponto de corte de 9. Desta forma, no presente estudo, o SRSI assume-se como capaz de discriminar a *performance* inerente a ambas as condições (Hipótese 5 comprovada), através dos pontos de corte preconizados pelos seus autores (Merten et al., 2016). Em consonância com o referido estão os resultados apurados em contexto internacional (e.g., Boskovic et al., 2019; Helvoort et al., 2019). De forma detalhada, Helvoort e colaboradores (2019) reportaram um aumento estatisticamente significativo [t(38)=11.16; p=<.001; d =1.79] dos pseudo-sintomas na condição T2 (simuladores instruídos),

comparativamente à T1 (honestidade): enquanto na condição T1 apenas 2 participantes ultrapassaram o ponto de corte de 6 e nenhum o de 9, na condição T2 36 dos 39 participantes excederam o ponto de corte de 6, tendo 31 obtido um resultado acima do ponto de corte de 9. Desta forma, este estudo concluiu igualmente que os referidos pontos de corte diferenciaram de forma funcional os resultados da escala de *Pseudo-sintomas*, de ambas as condições.

Uma vez tendo em consideração a elevada prevalência e os motivos que levam pessoas reclusas a simular e/ou exagerar sintomatologia (e.g., acesso a medicação; aborrecimento face ao processo de avaliação, etc.), a inserção de uma amostra de controlo (participantes não reclusos) permitiu testar a capacidade discriminativa do SRSI para diferenciar (respostas de) participantes reclusos de participantes não-reclusos. De facto, o instrumento em escrutínio discriminou ambos os grupos, na medida em que, globalmente, se verificaram pontuações estatisticamente mais elevadas ($p<.05$) na amostra prisional, comparativamente à amostra de controlo, sendo o tamanho das diferenças grande (*d* de Cohen >.80). Não obstante, é de sublinhar a dificuldade em homogeneizar variáveis como a idade e escolaridade (ou seja, a ausência de controlo das mesmas) de ambas as amostras, o que constitui uma limitação subjacente ao exame da Hipótese 6.

Para além da suprarreferida análise, são de expor evidências adicionais. Com o mencionado, o SRSI revelou ser capaz de discriminar as respostas dos indivíduos instruídos a simular sintomas, daqueles que não se esperava que evidenciassem tal conduta, considerando os pontos de corte aferidos pelos autores. Neste âmbito, calculámos (e analisámos qualitativamente) a proporção *ratio*, nomeadamente para os participantes cujas pontuações se revelaram acima do ponto de corte (>6 e/ou >9) na escala de *Pseudo-sintomas*. Com o referido, pretendemos auxiliar a identificação de tendências de simulação de psicopatologia, na medida em que um *ratio* inferior àquele que está associado ao ponto de corte selecionado poderia ser indicativo de distorções de resposta significativas, pelo que um *ratio* superior poderia sinalizar a presença de falsos positivos (Merten et al., 2016). Na presente amostra prisional, a proporção *ratio* foi tão sinalizadora dessa tendência quanto os pontos de corte de rastreio (>6) e de diagnóstico (>9), em ambas as condições. Identicamente, no estudo de Helvoort e colaboradores (2019), o índice de *ratio* convergiu a identificação de simulação pelo ponto de corte >6, realçando a conclusão de que o SRSI é razoavelmente eficaz na deteção de condutas de simulação ou exagero de sintomatologia. Por sua vez, analisámos as respostas aos itens do SRSI que medem a consistência de resposta (Merten et al., 2016), sobretudo nos casos em que a validade de sintomas dos participantes se revelou dúbia na condição "honesta" (com resultados na escala de *Pseudo-sintomas* superiores ao ponto de corte >6 e/ou >9). Na maioria dos casos (14 dos 18 participantes), a referida análise facultou uma convergência de informação, na medida em que identificou contradições na apresentação de sintomatologia por parte dos participantes, alertando para a potencial adoção de um estilo de resposta enviesado (cf. Anexo G).

Uma das potencialidades do presente estudo recai no uso do *design* de simulação análogo, incrementador de uma validade interna forte, na medida

em que permite a distribuição dos indivíduos por condições distintas, baseando-se em procedimentos que incluem instruções estandardizadas. Todavia, tal como fora anteriormente exposto, uma limitação emerge aquando do uso do referido *design*: a incerteza da generalização dos resultados obtidos com os simuladores experimentais para os que se inserem em contexto real (Rogers & Gillard, 2011). Foi neste contexto que se afigurou necessidade de, na presente investigação, considerar e refletir as recomendações acerca das instruções experimentais, da construção da *cover story* e do QPE.

No que concerne à construção da *cover story*, as recomendações refletiram-se na apresentação clara e compreensível do requerido; na introdução de um aviso acerca da existência de indicadores de avaliação de condutas de simulação; entre outros. Tal visou proporcionar aos participantes informações úteis acerca do funcionamento dos TVS e do modo como os sintomas deveriam ser exagerados e/ou simulados (de forma a não serem detetados por aqueles), para que a situação experimental se aproximasse à prática de *coaching*, verificada em circunstâncias reais.

O QPE afigurou-se pertinente para um controlo da manipulação da condição experimental. As questões 1 ("*Considera este cenário compreensível?*") e 5 ("*Fez o seu melhor para comunicar os sintomas de modo convincente e credível?*") permitiram avaliar o esforço depositado (90% dos indivíduos responderam afirmativamente à questão 5); avaliar a compreensão das instruções (98% respondeu "sim" à questão 1) e confirmar a recordação do requerido (posteriormente à introdução da *cover story*, proporcionaram-se condições para que, oralmente, o participante tivesse oportunidade de explicitar, por palavras suas, o que estava a ser requerido) – estratégias identificadas como úteis, por Rogers (2018c). Desta forma, apesar de o esforço e as respostas fornecidas, por parte dos participantes, não se percecionarem como contestáveis (no geral, consideraram ter sido fácil interpretar e agir consoante o requerido na *cover story* – questão 4), a conjugação dos dados obtidos através das respostas às questões 2 e 3 ("*Considera este cenário viável e realista?*" e "*Achou compreensível que o personagem simulasse sintomas?*") parece refletir que, embora a maioria dos participantes acredite que o cenário é tipicamente observado na comunidade ("*Muitos amigos meus já o fizeram!*"- sic), a mesma maioria não o perceciona justificável e/ou compreensível ("*Eu não o faria!*" - sic). Assim, esta postura pode ter influenciado os resultados, dada a dificuldade acrescida inerente ao facto de os participantes se perspetivarem a agir de acordo com a situação descrita (Merten et al., 2005), sendo igualmente de considerar a hipótese de a referida postura ter sido influenciada por outros estilos de resposta (e.g., *faking good*), cuja avaliação se revelou ausente na presente investigação.

Outras limitações do presente estudo e futuras recomendações de investigação merecem ser comentadas. Primeiramente, uma vez que a presente amostra é constituída por homens inseridos no contexto prisional, resta indagar se os resultados obtidos podem ser generalizados para outras amostras (e.g., com população feminina) ou grupos clínicos (e.g., reclusos internados em hospital psiquiátrico – com verdadeira psicopatologia). Para tal, seria relevante potenciar investigações nacionais complementares, que

tomassem em consideração o referido, de forma a analisar padrões de comportamentos de simulação inerentes a diferentes populações. Por sua vez, e tendo em conta a inserção num contexto prisional, alguns tipos de psicopatologia (ainda que não diagnosticada) poderão ter estado sobrerrepresentadas (e.g., patologia de personalidade – cf. Brazão et al., 2015) e, em consequência, influenciado a resposta ao protocolo de avaliação.

Quanto à extensão do protocolo de avaliação, optámos por mantê-la o mais breve possível, tendo em conta as características típicas da população prisional (e.g., diminuta escolaridade), de forma a minimizar potencial fadiga. Por esta razão, os instrumentos EVS-2 e SIMS foram aplicados, cada um, somente a metade da amostra. Por sua vez, não foi incluída uma medida de desejabilidade social (e.g., EDS-20), que teria sido pertinente para medir a tendência de *faking good* e outros estilos de resposta defensivos (e para corroborar, ou infirmar, a possibilidade anteriormente levantada – potencial intenção de *faking good*, no que concerne à validade de respostas tecidas ao QEP, sobretudo à sua 3ª questão), uma vez que os TVS se preveem incapazes de o fazer (Giger & Merten, 2019). Sem tal medida, a exclusão de participantes que evidenciaram um estilo de resposta de *fake good* não se concretizou no presente estudo, pelo que futuros estudos deverão incluí-la.

Outra limitação recai no facto de, no presente estudo, os participantes terem sido instruídos a simular sintomas, sem termos incluído um adicional "grupo conhecido" que exibisse, *per si*, uma validade de sintomas pobre (ou seja, que se percecionasse como tendo intenções reais e intrínsecas de simulação de sintomas). Desta forma, tal comprometeu a validade externa do presente estudo, pelo que, em futuras investigações seria pertinente concomitar ambas as tipologias de *design*, para que se potenciem resultados com validades interna e externa elevadas. Por sua vez, o uso de incentivos positivos (e.g., monetários) é um método tipicamente adotado no âmbito do *design* de simulação análogo, para que seja incrementada a envolvência ativa e o esforço do participante, mas tal não foi implementado no presente estudo. Por fim, como forma de (re)testar a sensibilidade à real psicopatologia por parte da escala de *Pseudo-sintomas*, seria relevante concretizar um estudo, em contexto nacional, que se assemelhasse ao concretizado por Helvoort e colaboradores (2019): ao integrar um grupo de pacientes-participantes (respondente na condição "honesta") com real psicopatologia, observar-se-ia se a escala *Pseudo-sintomas* os identificava, ou não, como simuladores.

VI - Conclusões

Uma vez que o diagnóstico de comportamentos de simulação e/ou de exagero de psicopatologia encerra em si um dos juízos clínicos que mais implicações pejorativas acarreta, a necessidade de se estabelecerem medidas precisas e capazes de os detetar tem sido um foco investigativo das últimas décadas. Tendo em consideração a prevalência da intencionalidade ou motivação para fabricar ou exagerar sintomas inerente ao contexto forense e prisional (mas não apenas; Kemp et al., 2008), torna-se premente a

administração de diversas medidas de validade de sintomas (TVS), que possam aliar-se, na tentativa de evitar a ocorrência do erro máximo (Tipo I) – a sinalização de indivíduos como simuladores, quando, na verdade, não recorreram a tal estilo de resposta negativo (Slick et al., 1999).

A presente investigação procurou contribuir para a prossecução da validação e exploração das potencialidades psicométricas do SRSI no contexto prisional português. O SRSI é um TVS recentemente construído que objetiva detetar a produção e/ou o exagero de sintomatologia, de forma a colmatar as limitações inerentes à maioria das medidas isoladas de validade de sintomas, sobretudo do SIMS – o instrumento mais amplamente utilizado neste âmbito avaliativo (Merten et al., 2016). Inversamente ao SIMS, o SRSI ostenta características que o destacam positivamente: apresenta uma escala que comporta sintomas tipicamente reportados (designados genuínos) em casos de litígio civil, não integrando na sua constituição apenas sintomas atípicos ou bizarros. Tal faz com que o SRSI, por um lado, detete mais facilmente simuladores sofisticados (com conhecimentos prévios acerca de quadros psicopatológicos) e, por outro, se torne mais resistente ao *coaching*.

De forma sucinta, prevê-se pertinente destacar as principais conclusões surgentes do presente estudo. No âmbito dos estudos de validade convergente, foi possível constatar relações fortes entre as pontuações da escala de *Pseudo-sintomas* (SRSI), e as verificadas no SIMS e na EVS-2, refletindo que pontuações elevadas na primeira se associam a resultados elevados no SIMS (resultados consonantes com os apurados em Domingues, 2019; Dwarkadas, 2018; Helvoort et al., 2019; Merten et al., 2016; Pinheiro, 2019) e na EVS-2 (idêntico ao verificado por Dwarkadas, 2018). Tal fornece suporte para confirmar a validade da escala de *Pseudo-sintomas* como um índice eficaz de relato distorcido de sintomatologia, o que permite concluir que o SRSI é um instrumento com capacidade para se afirmar como um complemento útil ao SIMS (Martin et al., 2015), respondendo, assim, à necessidade de recorrer a diversos TVS para fundar conclusões diagnósticas.

Não obstante o referido, verificaram-se correlações moderadas e fortes entre ambas a escalas do SRSI e os índices globais do BSI – medida de avaliação de psicopatologia real. Tal sugere que a escala de *Pseudo-sintomas* pode ser sensível à presença de psicopatologia genuína, o que constituiu uma limitação importante do instrumento (cf. no mesmo sentido: Boskovic et al., 2019; Dwarkadas, 2018; Pinheiro, 2019) – aspeto sinalizador de ausência de validade discriminante –, contrapondo os resultados observados por Helvoort e colaboradores (2019).

O presente estudo recorreu a um *design* de simulação análogo (DAS), o que proporcionou verificar, por meio da comparação direta entre respostas "honestas" e respostas "simuladas", a existência de diferenças estatisticamente significativas entre ambas as condições (condição "honesta" *versus* condição "simulação"). Neste sentido, é de assinalar o padrão que se tem verificado nesta tipologia de investigação, intra ou inter-sujeitos (e.g., Boskovic et al., 2019; Domingues, 2019; Helvoort et al., 2019): na condição "simulação", comparativamente à condição "honestidade", os indivíduos apresentam significativamente mais (pseudo) sintomas em ambas as escalas do SRSI.

Desta forma, no presente estudo, o instrumento em escrutínio demonstrou ser capaz de discriminar a *performance* inerente a ambas as condições, através dos pontos de corte recomendados pelos seus autores (Merten et al., 2016). Igualmente, o SRSI demonstrou boa validade discriminativa para diferenciar a *performance* dos participantes reclusos de participantes não-reclusos (amostra de controlo).

Em suma, é de destacar que, apesar da vincada limitação que se verificou (a escala *Pseudo-sintomas* demonstrou ser vulnerável a real psicopatologia – carência de validade discriminante), as restantes características psicométricas obtidas na presente investigação permitem concluir confiabilidade, no que concerne ao uso do SRSI como instrumento de avaliação da credibilidade do autorrelato. Ainda, ao basearmos a presente investigação empírica num *design* de simulação análogo, cremos que os resultados obtidos permitiram alcançar informações que contribuíram para o enriquecimento da investigação nacional e internacional, no âmbito da avaliação e deteção de condutas de simulação e de exagero de sintomas com o SRSI.

Bibliografia

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (Eds.). (2014). *Standards for educational and psychological testing*. American Educational Research Association.

American Psychiatric Association (2013). *Manual de diagnóstico e estatístico das perturbações mentais* (5a ed.). Climepsi Editores.

APA (1987). *Diagnostic and statistical manual of mental disorders* (3rd ed.). American Psychiatric Association.

Berry, D. T. R. & Nelson, N. W. (2010). DSM-5 and malingering: A modest proposal. *Psychological Injury and Law*, *3*(4), 295-303. http://dx.doi.org/10.1007/s12207-010-9087-7

Berry, D. T. R., Lamb, D. G., Wetter, M. W., Baer, R. A., & Widiger, T. A. (1994). Ethical considerations in research on coached malingering. *Psychological Assessment*, *6*(1), 16–17. https://doi.org/10.1037/1040-3590.6.1.16

Billiet, J. B., & McClendon, M. J. (2000). Modeling acquiescence in measurement models for two balanced sets of items. *Structural Equation Modeling*, *7*(4), 608-628. https://doi.org/10.1207/S15328007SEM0704

Bland, J., & Altman, D. (1997). Statistics notes: Cronbach's alpha. *British Medical Journal*, *31*(4), 572-572. https://doi.org/10.1136/bmj.314.7080.572

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. http://doi.org/10.1037/1040-3590.18.4.415

Boone, K. B. (2007). A reconsideration of the Slick et al. (1999) criteria for malingered neurocognitive dysfunction. In K. B. Boone (Ed.), *Assessment of feigned cognitive impairment: A Neuropsychological perspective* (pp. 29-49). The Guilford Press.

Boone, K. B. (Ed.) (2021). *Assessment of feigned cognitive impairment: A neuropsychological perspective* (2nd ed.). The Guilford Press.

Boskovic, I., Biermans, A. J., Merten, T., Jelicic, M., Hope, L., & Merckelbach, H. (2018). The Modified Stroop Task is susceptible to feigning: Stroop performance and symptom over-endorsement in feigned test anxiety. *Frontiers in Psychology*, *9*, 1195. http://doi.org/10.3389/fpsyg.2018.01195

Boskovic, I., Merckelbach, H., Merten, T., Hope, L., & Jelicic, M. (2019). The Self-Report Symptom Inventory as an instrument for detecting over reporting: An exploratory study with instructed simulators. *European Journal of Psychological Assessment*, *36*(5), 730-739. https://doi.org/10.1027/1015-5759/a000547

Boskovic, I., Merten, T., & Merckelbach, H. (2021). How plausible is the implausible? Students' plausibility and prevalence ratings of the self-report symptom inventory. *Psychological Injury and Law*, *14*, 127-

133. https://doi.org/10.1007/s12207-021-09409-x

Brazão, N., da Motta, C., Rijo, D., & Pinto-Gouveia, J. (2015). The prevalence of personality disorders in portuguese male prison inmates: Implications for penitentiary treatment. *Análise Psicológica*, *33*(3), 279–289. https://doi.org/10.14417/ap.975

Bush, S. S., Heibronner, R. L., & Ruff, R. M. (2014). Psychological assessment of symptom and performance validity, response bias, and malingering: Official position of the Association for scientific advancement in psychological injury and law. *Psychological Injury and Law, 7*(3), 197-205. https://doi.org/10.1007/s12207-014-9198-7.

Canavarro, M. C. (2007). Inventário de Sintomas Psicopatológicos (BSI): Uma revisão crítica dos estudos realizados em Portugal. In M. R. Simões, C. Machado, M. M. Gonçalves & L. S. Almeida (Coord.), *Avaliação psicológica: Instrumentos validados para a população portuguesa* (pp. 305-331). Quarteto Editora.

Cernovsky, Z., Mendonça, J. D., Ferrari, J. R., Sidhu, G., Velamoor, V., Mann, S., … Woodbury-Fariña, M, A. (2019). Content validity of the affective disorder subscale of the SIMS. *Archives of Psychiatry and Behavioral Sciences*, *2*(2), 33-39.

Cernovsky, Z., Mendonça, J. D., Oyewumi, L. K., Ferrari, J. R., 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.

Cima, M., Hollnack, S., Kremer, K., Knauer, E., Schellbach-Matties, R., Klein, B., & Merckelbach, H. (2003). Structured Inventory of Malingered Symptomatology: SIMS. *Der Nervenarzt*, *74*(11), 977–986. https://doi.org/10.1007/s00115-002-1438-5

Clark, L. A., & Watson, D. (2019). Constructing validity: New developments in creating objective measuring instruments. *Psychological Assessment*, *31*(12), 1412-1427. https://doi.org/10.1037/pas0000626

Cunha, C. (2011). *Escala de Validade de Sintomas – Versão Experimental* (EVS- VE): Estudos de "simulação" e validação em amostra da comunidade (Dissertação de mestrado). Faculdade de Psicologia e de Ciências da Educação da Universidade de Coimbra, Coimbra, Portugal.

Dandachi-FitzGerald, B., Duits, A. A., Leentjens, A. F. G., Verhey, F. R. J., & Ponds, W. H. M. (2020). Performance and symptom validity assessment in patients with apathy and cognitive impairment. *Journal of the International Neuropsychology Society*, *26*(3), 314-321. http://doi.org/10.1017/S1355617719001139

Derogatis, L. R., & Spencer, P. M. (1982). *Brief Symptom Inventory: Administration, scoring, and procedure manual*. Clinical Psychometric Research.

DeVellis, R. F. (2003). *Scale development: Theory and applications* (2nd ed.). Sage.

Domingues, M. V. (2019). *Self-Report Symptom Inventory* (SRSI): Validação

através de um estudo de simulação análogo (Dissertação de Mestrado). Faculdade de Psicologia e Ciências da Educação da Universidade de Coimbra, Coimbra, Portugal.

Dwarkadas, A. K. (2018). *Self-Report Symptom Inventory (SRSI):* adaptação, validação e normas com base em estudos na comunidade (Dissertação de mestrado). Faculdade de Psicologia e Ciências da Educação da Universidade de Coimbra, Coimbra, Portugal.

Easton, F. & Akehurst, L. (2011). Tools for the detection of lying and malingering in the medico-legal interview setting. *Medico-Legal Journal*, *79*(3), 103-108. http://doi.org/10.1258/mlj.2011.011023

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*(1), 33-42. https://doi.org/10.1080/00223890709336832

Ferguson, K. E. (2004). Detecting malingering in forensic neuropsychological evaluations in litigants with mild traumatic brain injury. In W. T. O'Donohue & E. R. Levensky (Eds.), *Handbook of forensic psychology* (pp. 301-314). Elsevier Academic Press.

Geurten, M., Meulemans, T., & Seron, X. (2018). Detecting over-reporting of symptoms: the French version of the Self-Report Symptom Inventory. *The Clinical Neuropsychologist*, *32*(1), 164-181. https://doi.org/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*(2), 5-13. https://doi.org/10.1024/1421-0185/a000218

Gillard, N. D. (2010). Methodological issues in malingering research: The use of simulation designs (Master thesis). University of North Texas, North Texas, United States of America.

Greiffenstein, F. M., Baker, J. W. & Gola, T. (1994). Validation of malingered amnesia measures with a large clinical sample. *Psychological Assessment*, *6*(3), 218-224. https://doi.org/10.1037/1040-3590.6.3.218

Heide, van, D., Boskovic, I., & Merckelbach, H. (2017). Standard symptom inventories for asylum seekers in a psychiatric hospital: Limited utility due to poor symptom validity. *Psychological Injury and Law*, *10*(4), 358–367. https://doi.org/10.1007/s12207-017-9302-x

Heilbronner, R. L., & Bush, S. S. (2017). Neuropsychological evaluations in disability determinations. In S. S. Bush, G. J. Demakis, & M. L. Rohling (Eds.), *APA handbook of forensic neuropsychology* (pp. 295–308). American Psychological Association. https://doi.org/10.1037/0000032-013

Heilbronner, R. L., Sweet, J. J., Morgan, J. E., Larrabee, G. J., Mills, 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.

https://doi.org/10.1080/13854040903155063

Heilbrun, K. (1992). The role of psychological testing in forensic assessment. *Law and Human Behavior*, *16*(3), 257-272. https://doi.org/10.1007/BF01044769

Helvoort, D. van, 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*, *33*(6), 1069-1082. https://doi.org/10.1080/13854046.2018.1559359

Horton, A. M. & Reynolds, C. R., (Eds.) (2021). *Detection of malingering during head injury litigation* (3rd ed.). Springer.

Huhnt, B. (2013). Stone cold manipulation or just an excuse: Psychopathy and blame attribution as possible explanations for faking symptoms in a forensic setting (Master thesis). Maastricht University, Maastricht, Netherlands.

Impelen, A. van, 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. https://doi.org/10.1080/13854046.2014.984763

Impelen, A. van, Merckelbach, H., Jelicic, M., Niesten, I. J. M., & Campo, J. (2017). Differentiating factitious from malingered symptomatology: The development of a psychometric approach. *Psychological Injury and Law*, *10*(4), 341-357. https://doi.org/10.1007/s12207-017-9301

Kemp, S., Coughlan, A. K., Rownottom, C., Wilkinson, K., Teggart, V., & Baker, G. (2008). The base rate of effort test failure in patients with medically unexplained symptoms. *Journal of Psychosomatic Research*, *65*(4), 319-325. https://doi.org/10.1016/j.jpsychores.2008.02.010

Kucharski, L. T., Falkenbach, D. M., Egan, S. S., & Duncan, S. (2006). Antisocial personality disorder and the malingering of psychiatric disorder: A study of criminal defendants. *International Journal of Forensic Mental Health*, *5*(2), 195–204. https://doi.org/10.1080/14999013.2006.10471243

Larrabee, G. (2008). Aggregation across multiple indicators improves the detection of malingering: Relationship to likelihood ratios. *The Clinical Neuropsychologist*, *22*(4), 666-679. https://doi.org/10.1080/13854040701494987

Larrabee, G. J. (2005). Assessment of malingering. In G. J. Larrabee (Ed.), *Forensic neuropsychology: A scientific approach* (pp. 115–158). Oxford University Press.

Larrabee, G. J. (2007). Identification of malingering by pattern analysis on neuropsychological tests. In G. J. Larrabee (Ed.), *Assessment of malingered neuropsychological deficits* (pp. 80-99). Oxford University Press.

Larrabee, G. J., Greiffenstein, M. F., Greve, K. W., & Bianchini, K. J. (2007). Refining diagnostic criteria for malingering. In G. J. Larrabee (Ed.),

Assessment of malingered neuropsychological deficits (pp. 334–371). Oxford University Press.

Marchi, B. De & Balboni, G. (2018). Detecting malingering mental illness in forensics: Known-group comparison and simulation design with MMPI-2, SIMS and NIM. *Peer J*, *6*(*2*), :e5259. https://doi.org/10.7717/peerj.5259

Marôco, J. (2014). *Análise estatística com o SPSS Statistics* (6th ed.). ReportNumber.

Martin, P. K., Schroeder, R. W., & Odland, A. P. (2015). Neuropsychologists' validity testing beliefs and practices: A survey of North American professionals. *The Clinical Neuropsychologist*, *29*(6), 741–776. https://doi.org/10.1080/13854046.2015.1087597

Martin, P. K., & Schroeder, R. W. (2020). Base rates of invalid test performance across clinical non-forensic contexts and settings. *Archives of Clinical Neuropsychology*, *35*(6), 717-725. https://doi.org/10.1093/arclin/acaa017

McCarter, R., Walton, N., Brooks, N., & Powell, G. (2009). Effort testing in contemporary UK neuropsychological practice. *The Clinical Neuropsychologist*, *23*(6), 1050-1066. https://doi.org/10.1080/13854040802665790

McDermott, B. E. & Sokolov, G. (2009). Malingering in a correctional system: The use of the Structured Interview of Reported Symptoms in a jail sample. *Behavioral Sciences and the Law*, *27*(5), 753-765. https://doi.org/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. https://doi.org/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*(3), 450–470. https://doi.org/10.1037/a0019216

Merckelbach, H. & Smith, G. P. (2003). Diagnostic accuracy of the Structured Inventory of Malingered Symptomatology (SIMS) in detecting malingering. *Archives of Clinical Neuropsychology*, *18*(2), 145-152. https://doi.org/10.1016/S0887-6177(01)00191-3

Merten, T. & Merckelbach, H. (2013). Symptom validity testing in somatoform and dissociative disorders: A critical review. *Psychological Injury and Law*, *6*(2), 122-137. https://doi.org/10.1007/s12207-013-9155-x

Merten, T., Dandachi-FitzGerald, B., Hall, V., Schmand, 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. https://doi.org/10.1016/S1130-5274(13)70014-8

Merten, T., Green, P., Henry, M., Blaskewitz, N., & Brockhaus, R. (2005). Analog validation of German-language symptom validity tests and the influence of coaching. *Archives of Clinical Neuropsychology*,

20(6), 719-726. https://doi.org/10.1016/j.acn.2005.04.004

Merten, T., Kaminski, A., & Pfeiffer, W. (2019). Prevalence of overreporting on symptom validity tests in a large sample of psychosomatic rehabilitation inpatients. *The Clinical Neuropsychologist*, *34*(5), 1004-1024. https://doi.org/10.1080/13854046.2019.1694073

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. Psyc*hological Injury and Law*, *9*, 102-111. https://doi.org/10.1007/s12207-016-9257-3

Mittenberg, W., Patton, C., Canyock, E. M., & Condit, D. C. (2002). Base rates of malingering and symptom exaggeration. *Journal of Clinical and Experimental Neuropsychology*, *24*(8), 1094-1102. https://doi.org/10.1076/jcen.24.8.1094.8379

Mota, M. (2008). Comportamentos de "simulação" e "esforço insuficiente": Estudos com a Structured Interview of Reported Symptoms (SIRS), o Structured Inventory of Malingering of Symptomatology (SIMS) e o Test of Memory Malingering (TOMM) numa amostra de reclusos (Dissertação de Mestrado). Faculdade de Psicologia e Ciências da Educação da Universidade de Coimbra, Coimbra, Portugal.

Nijdam-Jones, A., & Rosenfeld, B. (2017). Cross-cultural feigning assessment: A systematic review of feigning instruments used with linguistically, ethnically, and culturally diverse samples. *Psychological Assessment*, *29*(11), 1321–1336. https://doi.org/10.1037/pas0000438

Pankratz, L., & Erickson, R. C. (1990). Two views of malingering. *Clinical Neuropsychologist*, *4*(4), 379–389. https://doi.org/10.1080/13854049008401832

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

Rassin, E., Eerland, A., & Kuijpers, I. (2010). Let's find the evidence: An analogue study of confirmation bias in criminal investigations. *Journal of Investigative Psychology and Offender Profiling*, *7*(3), 231-246. https://doi.org/10.1002/jip.126

Resnick, P. J. & Knoll, J. L. (2018). Malingered psychosis. In R. Rogers & S. D. Bender (Eds.), *Clinical assessment of malingering and deception* (4th ed., pp. 98-121). The Guilford Press.

Resnick, P. J. (1997). Malingered psychosis. In R. Rogers (Ed.), *Clinical assessment of malingering and deception* (1st ed., pp. 47-67). The Guilford Press.

Rogers, R. & Bender, S. (2012). Evaluation of malingering and related response syles. In I. B. Weiner (Ed.), *Handbook of psychology – Vol.11: Forensic psychology* (2nd ed., pp. 517-540). Wiley.

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

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

Rogers, R. (2018a). An introduction to response styles. In R. Rogers & S. D. Bender (Eds.), *Clinical assessment of malingering and deception* (4th ed, pp. 3-17). *The Guilford Press.*

Rogers, R. (2018b). Detection strategies for malingering and defensiveness. In R. Rogers & S. D. Bender (Eds.), *Clinical assessment of malingering and deception* (4th ed., pp. 18-41). The Guilford Press.

Rogers, R. (2018c). Researching response styles. In R. Rogers & S. D. Bender (Eds.), *Clinical assessment of malingering and deception* (4th ed., pp. 592-614). The Guilford Press.

Rogers, R., & Gillard, N. D. (2011). Research methods for the assessment of malingering. In B. Rosenfeld & S. D. Penrod (Eds.), *Research methods in forensic psychology* (pp. 174–188). John Wiley & Sons.

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. https://doi.org/10.1207/s15327752jpa6702_2

Rogers, R., Jackson, R. L., Sewell, K. W., & Harrison, K. S. (2004). An examination of the ECST-R as a screen for feigned incompetency to stand trial. *Psychological Assessment*, *16*(2), 139–145. https://doi.org/10.1037/1040-3590.16.2.139

Rogers, R., Robinson, E. V., & Gillard, N. D. (2014). The SIMS screen for feigned mental disorders: The development of detection-based scales. *Behavioral Sciences & the Law*, *32*(4), 455-466. https://doi.org/10.1002/bsl.2131.

Ruiz, M. A., Drake, E. B., Glass, A., Marcotte, D., & van Gorp, W. G. (2002). Trying to beat the system: Misuse of the internet to assist in avoiding the detection of psychological symptom dissimulation. *Professional Psychology: Research and Practice*, *33*(3), 294–299. https://doi.org/10.1037/0735-7028.33.3.294

Savitsky, K., & Gilovich, T. (2003). The illusion of transparency and the alleviation of speech anxiety. *Journal of Experimental Social Psychology*, *39*(6), 618–625. https://doi.org/10.1016/S0022-1031(03)00056-8

Seron, X. (2014). Lying in neuropsychology. *Neurophysiologie Clinique/ Clinical Neurophysiology*, *44*(4), 389-403. http://dx.doi.org/10.1016/j.neucli.2014.04.002

Sherman, E. M. S., Slick, D. J., & Iverson, G. L. (2020). Multidimensional malingering criteria for neuropsychological assessment: A 20-year update of the malingered neuropsychological dysfunction criteria. *Archives of Clinical Neuropsychology*, *35*(6), 735-764. http://dx.doi.org/10.1093/arclin/acaa019

Simões, M. R. (2006). Testes de validade de sintomas na avaliação de comportamentos de simulação. In A. C. Fonseca, M. R. Simões, M. C. T. Simões & M. S. Pinho, (Eds.), *Psicologia forense* (pp. 280-

309). 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). Pactor.

Simões, M. R., Almeida, L. S., & Gonçalves, M. M. (2017). Instrumentos de avaliação em psicologia forense: Contributos da investigação para a prática profissional. In M. R. Simões, L. S. Almeida, & M. M. Gonçalves (Coord.), *Psicologia forense: Instrumentos de avaliação* (pp. 1-22). Pactor.

Simões, M. R., Almiro, P. A., Mota, M., Oliveira, C. S., 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 (Coord.), *Psicologia forense: Instrumentos de avaliação* (pp. 397-415). Pactor.

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 (Eds.), *Psicologia forense: Instrumentos de avaliação* (pp. 417-435). Pactor.

Simões, M. R., Sousa, L. B., Fonseca, M. S., Garcia, S. G., 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., Douro, 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 (Coord.), *Psicologia forense: Instrumentos de avaliação* (pp. 353-375). Pactor.

Simões, M.R. & Ferraz, M. (2021). Estilos de resposta: Validade, verdade e risco de arbitrariedade em Avaliação Neuropsicológica. In Laura Alho & Mauro Paulino (Coords.), *Psicologia do testemunho* (pp. 107-137). Pactor.

Slick, D., Sherman, E., & Iverson, G. L. (1999). Diagnostic criteria for malingered neurocognitive disfunction: Proposed standards for clinical practice and research. *The Clinical Neuropsychologist, 13*(4), 545-561. http://doi.org/10.1076/1385-4046(199911)13:04;1-Y;FT5

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*(2), 183-189.

Springman, R. E. & Vandenberg, B. R. (2009). The effects of teste strategy coaching on measures of competency to stand in trial. *Journal of Forensic Psychology Practice*, *9*(3), 179-198. http://doi.org/10.1080/15228930902809856

Steffan, J. S. & Morgan, R. (2008). Diagnostic accuracy of the MMPI-2 Malingering Index in the detection of malingering among inmates. Journal of Personality Assessment, *90*(4), 392-398.

https://doi.org/10.1080/00223890802108204

Suhr, J. A. & Gunstad, J. (2007). Coaching and malingering: A review. In G. J. Larrabee (Ed.), *Assessment of malingered neuropsychological deficits* (pp. 287-311). Oxford University Press.

Tombaugh, T. M. (1996). *Test of Memory Malingering (TOMM)*. Multi-Health Systems.

Tracy, D. K. & Rix, K. J. B. (2017). Malingering mental disorders: Clinical assessment. *Advances in Psychiatric Treatment, 23*(1), 27-35. http://doi.org/10.1192/apt.bp.116.015958

Velsor, S.F., Rogers, R., Donnelly, J.W., & Tazi, K. (2021). Assessment of Factitious Psychological Presentations (FPP): An overlooked response style in forensic practice. *Psychology Injury and Law, 14*, 201–212. https://doi.org/10.1007/s12207-021-09413-1

Victor, T. L. & Abeles, N. (2004). Coaching clients to take psychological and neuropsychological tests: A clash of ethical obligations. *Professional Psychology: Research and Practice, 35*(4), 373-379. http://doi.org/10.1037/0735-7028.35.4.373

Viglione, D. J., & Giromini, L. (2020). *Inventory of Problems–29: Professional manual*. OH: IOP-Test, LLC.

Viglione, D. J., Giromini, L., & Landis, P. (2017). The development of the Inventory of Problems–29: A brief self-administered measure for discriminating bona fde from feigned psychiatric and cognitive complaints. *Journal of Personality Assessment, 99*(5), 534–544. https://doi.org/10.1080/00223891.2016.1233882

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

Vitacco, M. J., 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. https://doi.org/10.1007/s10979-006-9062-8

Wall, T. D., Wygant, D. B., & Gallagher, R. W. (2015). Identifying overreporting in a correctional setting: Utility of the MMPI-2 restructured form validity scales. *Criminal Justice and Behavior, 42*(6), 610-622. https://doi.org/10.1177/0093854814556881

Wechsler, D. (2008). *Escala de Inteligência de Wechsler para Adultos-Terceira Edição (WAIS-III)*. Cegoc.

Widows, M. R., & Smith, G. P. (2005). *Structured Inventory of Malingered Symptomatology: Professional manual*. 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. https://doi.org/10.1093/arclin/acp110

Young, G. (2014). Assessing psychological injuries and malingering: Disability and report writing. In G. Young (Ed.), *Malingering, feigning, and response bias in psychiatric/psychological injury: Implications for practice and court* (pp. 361-384). Springer Science & Business Media.

Young, G. (2021). The call for aid (cry for help) in Psychological Injury and Law: Reinterpretation, mechanisms, and a call for research. *Psychological Injury and Law*, *14*, 185-200. https://doi.org/10.1007/s12207-021-09414-0

Young, G., Foote, W. E., Kerig, P. K., Mailis, A., Brovko, J., Kohutis, E. A., McCall, S., Hapidou, E. G., Fokas, K. F., & Goodman-Delahunty, J. (2020). Introducing psychological injury and law. *Psychological Injury and Law*, *13*, 452–463. https://doi.org/10.1007/s12207-020-09396-5-

Anexos

Anexo A

Quadro 1

Estudos com o SRSI em contexto prisional e na população portuguesa em geral.

Estudo e amostra	Objetivos	Conclusões relevantes
Huhn (2013) 65 Reclusos masculinos de uma prisão juvenil.	- Testar a aplicabilidade do SRSI num contexto criminal e forense.	• Correlação moderada entre a escala de Pseudo-sintomas (SRSI) e o SIMS (r= .49; p <0.05). O autor justifica a referida associação moderada com a passagem alargada de tempo (2 ou 3 semanas de intervalo) entre a primeira (aplicação do SRSI) e a segunda (aplicação do SIMS) sessão. • A correlação entre o SIMS e a escala de Sintomas Genuínos (SRSI) foi forte (r= .51; p <0.05).
Domingues (2019) 70 Estudantes universitários.	- Tradução e adaptação do SRSI para português europeu. - Validação do SRSI em contexto universitário, através de um *design* de simulação análogo (DSA).	• Incremento de sintomas reportados da primeira condição ("honesta") para a segunda ("de simulação"), expressando um notório empenho na condição de exagero e na simulação de sintomas, por parte dos participantes. • Média de *Sintomas Genuínos* mais elevada do que *Pseudo-sintomas* na condição "honesta" – expectável –, mas também na condição de "simulação" – relação considerada não expectável pela autora. • Inesperadamente, a correlação verificada entre as escalas *Sintomas Genuínos* e *Pseudo-sintomas* do SRSI e o SIMS revelaram-se muito fortes (r= .756; p <0.05 e r= .801; p <0.05 , respetivamente)
Dwarkadas (2018) 282 Indivíduos (estudantes universitários e/ou indivíduos adultos não estudantes).	- Tradução e adaptação do SRSI para português europeu. - Início do processo de validação do instrumento para a população portuguesa.	• Correlações problemáticas e não expectáveis, particularmente elevadas entre as pontuações da escala de *Sintomas Genuínos* (SRSI) e os resultados dos restantes TVS (SIMS e EVS-2). • Os resultados nos instrumentos de avaliação de psicopatologia revelaram correlações problemáticas com ambas as escalas do SRSI, podendo tal refletir sensibilidade à psicopatologia genuína, o que contraria dados provenientes da literatura internacional (e.g., Helvoort e colaboradores, 2019).

Estudo e amostra	Objetivos	Conclusões relevantes
Pinheiro (2019) 85 Reclusos masculinos.	-Tradução e adaptação para português europeu o SRSI. -Iniciar o processo de validação do SRSI em contexto prisional.	• Tal como em Dwarkadas (2018), foram encontradas neste estudo correlações problemáticas: correlações fortes e moderadas entre ambas as (sub)escalas do SRSI e o resultado total dos diversos instrumentos de avaliação de psicopatologia (BDI-II, STAI e BPI), sugerindo que o SRSI pode ser sensível à psicopatologia real (limitação). • Identificados itens controversos para a população prisional (ou seja, que parecem não considerar as características da população em causa – e.g., estar em situação de reclusão, reduzida escolaridade), tanto no que concerne ao SRSI, como ao SIMS. • A **extensão do protocolo** foi identificada como um fator limitador, já que potenciou a fadiga nos participantes, dificultando a sua aplicação, elementos particularmente notórios em contexto prisional (cujos participantes têm baixa escolaridade).
Helvoort et al. (2019) 40 Reclusos internados em hospital psiquiátrico.	-Testar se a escala de *Pseudo-sintomas* (SRSI) é sensível a psicopatologia genuína e avaliar a sua validade discriminante num estudo cuja amostra foi instruída a simular.	• Comparativamente ao SIMS, a escala de *Pseudo-sintomas* (SRSI) parece relativamente imune à presença de psicopatologia: os pacientes-participantes (com psicopatologia) apresentaram um número moderado de sintomas ditos genuínos (escala de *Sintomas Genuínos*), mas poucos pseudo-sintomas (escala de *Pseudo-sintomas*). • A percentagem de falsos positivos foi de 5% para o ponto de corte >6, não havendo registo de falsos positivos para o ponto de corte >9. É válida a utilização em contexto prisional de um instrumento que não produz falsos positivos em pacientes com efetiva psicopatologia.

Anexo B

Quadro 2

Conclusões relevantes de investigações empiricas (cuja metodologia assentou num DSA) com TVS, em contexto prisional.

Estudo e amostra	*Design* e TVS	Conclusões relevantes
Rogers et al. (1996) 53 Reclusos pacientes psiquiátricos.	- *Design* de Simulação Análogo (intra-sujeitos). - MMPI-A; SIRS; SIMS.	• A *Structured Interview of Reported Symptoms* (SIRS) tem uma probabilidade extremamente reduzida de identificar respondentes honestos como simuladores. • Como instrumento de rastreio, o SIMS provou ser moderadamente eficaz na identificação de simuladores.
Rogers et al. (2004) 96 Reclusos inseridos em zona comum; 56 Pacientes psiquiátricos.	- *Design* de Simulação Análogo (intra-sujeitos). - ECST-R; M-FAST; SIRS.	• Os pontos de corte estabelecidos para a escala ATP (*Atypical Presentation*) do *Evaluation of Competency to Stand Trial—Revised* (ECST-R) geraram poucos falsos negativos e apresentam estimativas robustas de sensibilidade. A ATP parece ser uma medida de rastreio de simulação razoável para uso em avaliações psicológicas de competência para ir a julgamento.
Edens et al. (2007) 60 Reclusos inseridos em zona comum; 30 Reclusos pacientes psiquiátricos; 26 Reclusos suspeitos de simulação de sintomas.	- *Design* de Simulação Análogo e Comparação de Grupos Conhecidos. - PAI (NIM, MAL e RDF); SIMS; SIRS.	• Embora os TVS usados tenham tido um bom desempenho nas amostras não psiquiátricas, a precisão da classificação foi mais reduzida na tentativa de diferenciar pacientes psiquiátricos de reclusos suspeitos de simulação de sintomas. Apenas 2 indicadores de PAI (*Malingering Index* [MAL] e *Rogers Discriminant Function* [RDF]) sido eficazes em discriminá-los significativamente.
Steffan & Morgan (2008) 45 Reclusos inseridos em zona comum; 46 Reclusos pacientes psiquiátricos.	- *Design* de Simulação Análogo. - MMPI-2.	• O índice M-DFI (*Malingering Discriminant Function Index* [MMPI-2]), desenhado para detetar simuladores instruídos, não evidenciou validade em amostras de reclusos com suspeita de psicopatologia simulada ou genuína.

Estudo e amostra	*Design* e TVS	Conclusões relevantes
Wall et al. (2015) 196 Reclusos inseridos em zona comum; 73 Reclusos pacientes psiquiátricos.	- *Design* de Simulação Análogo. - MMPI-2-RF.	• Os resultados indicaram tamanhos de efeito grandes (e.g., *d* de Cohen = 1.68) inerentes às diferenças verificadas entre os grupos de simulação e de controlo, para todas as escalas de validade do MMPI-2-RF (*Restructured Form*).
Marchi & Balboni (2018) 71 Reclusos inseridos em zona comum; 57 Reclusos pacientes psiquiátricos.	- *Design* de Simulação Análogo (DSA) e Comparação de Grupos Conhecidos. - MMPI-2 (Fc e Fp); SIMS; PAI (NIM).	• Todos os TVS – as escalas do MMPI-2 (Fc [*Criminal Offender Infrequency*] e Fp [*Infrequency Psychopatology*]), o SIMS e a escala NIM (*Negative Impression Management*; PAI) – foram eficazes na deteção de simuladores no DSA. • O SIMS discrimina os reclusos simuladores dos indivíduos pertencentes ao grupo de controlo (reclusos inseridos em zona comum) no *design* de Comparação de Grupos Conhecidos. Não obstante, a baixa especificidade do do SIMS limita a sua utilização como uma medida única de avaliação da simulação no contexto prisional.
Helvoort et al. (2019) 40 Reclusos pacientes psiquiátricos.	- *Design* de Simulação Análogo (intra-sujeitos). - SRSI.	• É válido o uso do SRSI em contexto prisional, uma vez que os seus resultados não produzem falsos positivos em pacientes com pacientes com efetiva psicopatologia.

Anexo C

Consentimento informado – participação em estudo científico

Este estudo decorre no âmbito da dissertação de Mestrado em Psicologia Forense da Faculdade de Psicologia e Ciências da Educação da Universidade de Coimbra (FPCE-UC), sob autoria da aluna Cláudia Venâncio, atualmente a frequentar o 5º ano do Mestrado Integrado em Psicologia. A presente investigação é orientada pelo Professor Doutor Mário R. Simões (Professor Catedrático da FPCE-UC).

O estudo tem por **objetivo** dar continuidade ao processo de **validação de um instrumento de avaliação psicológica para o contexto prisional português** (cf. estudo de Pinheiro, 2019). Trata-se de um questionário que contém frases que exprimem queixas ou dificuldades que as pessoas podem sentir no seu dia-a-dia.

Para que consigamos atingir tal objetivo, solicitamos a sua colaboração através da resposta a um conjunto de testes psicológicos, incluindo o questionário acima referido. A sua **participação é voluntária**, mas **muito importante para a realização deste estudo**.

Todas as respostas dadas pelo participante são **confidenciais**, analisadas de forma agregada (e não individualizada) e usadas apenas para fins estatísticos e científicos. É **assegurado** o **anonimato do participante**.

Enquanto participante, **tem o direito e a oportunidade de colocar as questões que considerar pertinentes, as quais serão respondidas de forma a esclarecê-lo**. A aluna e investigadora está **disponível para qualquer esclarecimento acerca do estudo**.

Pode optar por não querer dar continuidade à sua participação, pelo que pode desistir da sua participação a qualquer momento.

Muito obrigada pela sua disponibilidade e colaboração.

Eu, _____,

concordo em participar voluntariamente no presente estudo.

Adicionalmente, declaro que, antes do início da avaliação, me foram explicados os objetivos e o protocolo de avaliação. Foi-me assegurada a confidencialidade e o anonimato dos dados. Fui igualmente informado que, em qualquer momento, poderei desistir de colaborar no presente estudo.

Anexo D

FACULDADE DE PSICOLOGIA E DE CIÊNCIAS DA EDUCAÇÃO UNIVERSIDADE D COIMBRA

Questionário de dados sociodemográficos

IDENTIFICAÇÃO

Idade: ___ anos

Nacionalidade: _____

Estado civil: Solteiro ☐ Casado/ união de facto ☐ Divorciado ☐ Viúvo ☐

HABILITAÇÕES LITERÁRIAS

Anos de escolaridade atuais (exemplo: se tiver concluído o 9ºano, responda 9): _____

Tem dificuldades na compreensão da língua portuguesa? Não ☐ Sim ☐

SAÚDE

Apresenta algum diagnóstico clínico na área da saúde mental?

Não ☐ Sim ☐

Anexo E

Cenário experimental

Considere a seguinte situação:

Imagine que é um mecânico que trabalhou muitos anos numa oficina, mantendo-se motivado para o trabalho porque adorava a sua profissão. Numa tarde em que estava ao serviço, um cliente dirige-se à oficina e acusa-o de lhe ter **roubado peças do seu carro**, avisando-o de que **instaurou um processo judicial contra si**.

Imagine que este evento teve um grande impacto em si nos dias e semanas seguintes, causando-lhe um grande sofrimento: **ansiedade, pesadelos, pensamentos intrusivos ("*o que irão pensar de mim?*"), cansaço físico e psicológico (desconcentração)**. Ainda devido à referida situação, o seu patrão foi arrogante para si, não tendo em consideração o excelente profissional que sempre foi, nem o impacto que a situação teve em si, o que o deixou **mais ansioso e revoltado**.

Em consequência, sente que não consegue continuar a trabalhar na oficina, devido à pressão do processo judicial (e.g. prestar declarações junto das entidades judiciais) e ao sofrimento que a situação lhe trouxe, pelo que **decide pedir baixa médica**. Para tal, consulta o seu médico de família, que o encaminha para uma Consulta de Psicologia. Nesta consulta, o mecânico irá responder a testes psicológicos cujos resultados ajudarão a determinar se necessita, ou não, da referida baixa médica.

O mecânico acreditava que, **se <u>exagerasse</u> os sintomas sentidos ou <u>fingisse</u> outros (e.g. dores fortes no corpo e na cabeça; sentimentos depressivos; sem motivação e sem interesse por nada; memória diminuída**) na avaliação, **conseguiria obter a baixa médica com maior facilidade**. Assim, imagine que **decidiu fazê-lo**.

De seguida, **responda ao seguinte teste psicológico como se fosse o referido mecânico, apresentando os sintomas físicos e mentais acima referidos** (verdadeiros ou não) **de forma convincente**. Assim, tenha em atenção os sintomas que seleciona, **não esquecendo o objetivo que deseja alcançar com o exagero ou o fingimento dos seus problemas: receber baixa médica**.

Para que não seja descoberto, lembre-se que **não deve mencionar sintomas que possam parecer muito improváveis ou raros**, já que, ao fazê-lo, poderá ser mais facilmente identificado como **alguém que não está a ser honesto** e, portanto, não receberá a baixa médica.

Anexo F

Questionário Pós-experimental

1. **Considera este cenário compreensível** (ao nível da compreensão do que está a ser pedido)**?**

   Sim ☐ Não ☐ Moderadamente ☐

2. **Considera este cenário viável e realista** (com probabilidade de acontecer se estivesse no lugar deste personagem)?

   Sim ☐ Não ☐ Moderadamente ☐

3. **Achou compreensível que o personagem principal simulasse sintomas?**

   Sim ☐ Não ☐ Moderadamente ☐

4. **Foi fácil para si imaginar que era o personagem principal?**

   Sim ☐ Não ☐ Moderadamente ☐

5. **Fez o seu melhor para comunicar os sintomas de modo convincente e credível?**

   Sim ☐ Não ☐ Moderadamente ☐

Anexo G

Num primeiro momento, verifica-se a pontuação total da escala de *Pseudo-sintomas* do SRSI, de forma a identificar as pontuações superiores ao ponto de corte previamente selecionado[6]. Se for utilizado um dos pontos de corte mais baixos (e.g., >6) e, em consequência, o examinando tiver obtido uma pontuação acima do mesmo, é pertinente calcular a proporção *ratio*[7].

Qualitativamente, um *ratio* inferior àquele que está associado ao ponto de corte selecionado é indicativo de distorções de resposta significativas. Por sua vez, um *ratio* superior poderá evidenciar a presença de falsos positivos, sobretudo aquando do uso do *liberal cutoff* (Tabela 1).

Tabela 1

Pontos de corte propostos por Merten et al., 2016.

Pontos de corte para o Total de Pseudo-sintomas	Sensibilidade	Especificidade	*Ratio*	Pontos de corte selecionados
>4	.895	.828	5.20	*Liberal cutoff*
>6	.829	.911	9.31	*Proposed cutoff*
>9	.618	.955	13.73	*Cutoff to minimize false positives*
<15	.329	.990	32.90	*Most rigorous cutoff*

A título de exemplo, no âmbito da condição "honesta", para os participantes que obtiveram uma pontuação acima do ponto de corte de rastreio (>6) e/ou de diagnóstico (>9), a validade dos sintomas apresentados revelou-se duvidosa. Nesse sentido, verificaram-se os itens que medem a consistência de resposta do SRSI (cf. Tabela 2), que deveriam ser assinalados como falsos (para que a consistência não se revelasse dúbia). Se algum dos mesmos foi assinalado como verdadeiro, estaremos, segundo o racional do SRSI (Merten er al., 2016), perante contradições na apresentação de sintomas por parte do examinando-participante, o que, por sua vez, dá solidez à identificação, por parte do instrumento, de uma conduta de simulação de sintomatologia, tendo em conta os pontos de corte identificados.

[6] A seleção dos pontos de corte teve em consideração o referido por Iverson (2007), que refere que devemos selecionar aquele que apresenta um nível de especificidade mais elevado que o nível se sensibilidade, ponteciando, dessa forma, a redução de falsos positivos. Desta forma, analisámos os nossos dados tendo em consideração o ponto de corte de rastreio (>6) e/ou de diagnóstico (>9).

[7] *Ratio* = (total de *Sintomas Genuínos*) / (total de *Pseudo-sintomas* + 0.1).

Tabela 2

Valores da proporção ratio e dos itens que medem a consistência de resposta (SRSI), para os participantes cujas pontuações se situaram acima dos pontos de corte de rastreio (>6) e/ou de diagnóstico (>9), na escala Pseudo-sintomas (condição "honesta").

Participante	Número de pseudo-sintomas assinalados	*Ratio*	Consistência de resposta				
			Item 8	Item 29	Item 49	Item 75	Item 95
1	16	2.609	V	F	F	F	V
7	14	1.702	F	V	V	V	V
9	9	3.407	F	F	V	F	F
11	9	2.967	V	V	V	F	V
23	10	3.465	F	F	F	F	F
30	13	2.443	F	F	F	V	F
36	28	1.459	V	V	F	F	F
38	11	2.793	V	V	F	F	F
48	29	1.340	V	V	V	V	V
57	8	2.963	V	V	F	V	V
61	9	4.286	F	F	F	F	F
63	18	2.044	V	V	V	F	V
68	10	3.069	F	F	V	V	V
79	24	1.452	V	V	F	V	V
81	8	2.840	F	F	V	F	F
87	20	1.891	F	F	F	F	F
92	7	3.944	V	V	V	F	V
94	20	2.090	F	F	F	F	F

No que concerne à proporção *ratio*, todos os valores foram inferiores ao valor de ratio associado ao ponto de corte selecionado (tanto para o ponto de corte >6, como para os casos que ultrapassaram o ponto de corte >9), o que converge a identificação, por parte do instrumento, de distorções de resposta significativas (neste caso, de exagero/simulação de sintomatologia).

Por sua vez, a maioria dos participantes cuja pontuação na escala *Pseudo-sintomas* excedeu o ponto de corte de rastreio (à exceção dos participantes 23, 61, 87 e 94 – cf. Tabela 2) assinalou como verdadeiro pelo menos um dos itens que medem a consistência de resposta. Tal é indicativo de uma validade de autorrelato dúbia e, em consequência, de uma potencial adoção de um estilo de resposta negativo (e.g., exagero/ simulação de sintomatologia), dando solidez à identificação, por parte do instrumento, de uma conduta de simulação de sintomatologia, tendo em conta os pontos de corte identificados.