

Aplicação do *Rasch Rating Scale Model* ao Inventário de Sintomas Obsessivo-Compulsivos

Joana Rosa¹, Ana R. Delgado², Gerardo Prieto³,
Américo Baptista⁴ e Paulo Lopes⁵

Application of the *Rasch Rating Scale Model* to the Obsessive-Compulsive Symptoms Inventory

Abstract

The purpose of this study was to analyse, using the Rasch Rating Scale Model, an instrument of obsessive-compulsive symptoms (ISOC), presented in 2008 by Rosa and Baptista, that consists of 105 items of characteristic symptoms of obsessive-compulsive disorder. Two hundred and twenty subjects (87 diagnosed with OCD and 133 without this disorder) undertook the ISOC (Rosa & Baptista, 2008) and the SCL-90-R (Derogatis, 1977). It was confirmed that the response categories were adequate, although 27 items had to be removed since they showed differential functioning between groups (with OCD/ without OCD) and/or sex. The final version of the inventory consists of 78 items with five response categories that have met the criteria set (Linacre, 2002) and shows adequate psychometric qualities. Thus, the results obtained indicate that the ISOC is a reliable, valid and useful measure, useful as a screening instrument to assess obsessive-compulsive symptoms.

Keywords: obsessive-compulsive symptoms; Rasch model; response categories; psychometric study

1 Universidade Lusófona de Humanidades e Tecnologias, Lisboa, Portugal. Email: joanarosa@yahoo.com

2 Universidad de Salamanca. Email: adelgado@usal.es

3 Universidad de Salamanca, Email: gprieto@usal.es

4 Universidade Lusófona. Email: americo.baptista@sapo.pt

5 Universidade Lusófona. Email: pjflopes@gmail.com

Resumo

O objetivo deste estudo foi analisar, com o Modelo de Rasch, um inventário de sintomas obsessivo-compulsivo (ISOC), trabalhado por Rosa e Baptista, em 2008, que consiste de 105 itens de sintomas característicos da perturbação obsessivo-compulsiva (POC). Duzentos e vinte pacientes (87 com diagnóstico de POC e 133 sem esta perturbação) preencheram o ISOC (Rosa & Baptista, 2008) e a SCL-90-R (Derogatis, 1977). Verificou-se que as categorias de resposta se revelaram adequadas, embora 27 itens tenham mostrado funcionamento diferencial entre grupos (com POC/sem POC) e entre sexos, pelo que foram excluídos. A versão final do inventário ficou composta por 78 itens com cinco categorias de resposta que cumpriram os critérios estabelecidos (Linacre, 2002) e apresentaram qualidades psicométricas adequadas. Assim, os resultados evidenciaram que o ISOC é uma medida válida, confiável e útil como instrumento de *screening* para avaliar os sintomas obsessivo-compulsivos.

Palavras-chave: sintomas obsessivo-compulsivo; modelo de Rasch; categorias de resposta; estudo psicométrico

INTRODUÇÃO

A Perturbação Obsessivo-Compulsiva (POC) é uma desordem que envolve obsessões e/ou compulsões recorrentes e persistentes que consomem tempo e/ou causam grande dificuldade na vida diária (American Psychiatric Association - APA, 2013). As obsessões são pensamentos, imagens ou impulsos vivenciados como intrusivos e repetitivos e as compulsões são comportamentos ou atos mentais utilizados para tentar reduzir a ansiedade causada pela experiência das obsessões. A variedade de sintomas que a POC apresenta, a sua estranheza e o tempo gasto em pensamentos e/ou comportamentos estereotipados pode conduzir a graves consequências na qualidade de vida do paciente, bem como na da família e do círculo social (Franklin & Foa, 2011).

Em relação aos subtipos de POC, tradicionalmente a sua caracterização é efetuada com base nos sintomas prevalentes que, por norma, surgem no decorrer dos estudos realizados com as medidas desenvolvidas para a sua avaliação. Numa revisão efetuada por McKay e colaboradores (2004) encontram-se sintomas relacionados com a lavagem (medo da contaminação), verificação, armazenamento, simetria, ordenação, obsessões sexuais, religiosas ou somáticas.

Num estudo mais recente (Fullana et al., 2010), a prevalência das dimensões de sintomas obsessivo-compulsivos, as suas correlações com variáveis sociodemo-

gráficas e a comorbilidade foi investigada numa subamostra aleatória (N = 2804), a partir de um estudo transversal da população adulta geral em seis países europeus. Verificou-se que a prevalência, ao longo da vida, de qualquer dimensão de sintomas de POC foi de 13%. A dimensão Causar Dano/Verificação revelou-se a mais comum (8%), seguida pelas Obsessões somáticas (5%) e Simetria/Ordenação (3%). As mulheres eram mais propensas a ter sintomas de obsessões somáticas e de contaminação/limpeza.

Quanto à avaliação da POC, vários métodos têm sido utilizados, tais como questionários, entrevistas e testes comportamentais, cada um focando as manifestações da POC que foram consideradas as mais relevantes para os seus autores. A divisão clássica entre verificação e contaminação é comum, mas muitas outras dimensões ou facetas de conteúdos são refletidos nas medidas de autorrelato em vigor, como o *Leyton Obsessional Inventory* (LOI) (Cooper, 1970), o *Symptom Checklist-Revised-90* (SCL-90-R) (Derogatis, 1977), o *Maudsley Obsessive Compulsive Inventory* (MOCI) (Hodgson & Rachman, 1977), a *Compulsive Activity Checklist* (CAC) (Steketee & Freund, 1993), o *Inventário de Pádua* (PI_WSUR), (Burns, Keortge, Formea, & Sternberger, 1996), o *Vancouver Obsessive Compulsive Inventory* (Thordarson et al., 2004), o *Obsessive-Compulsive Inventory* (Foa, Kozak, Salkovskis, Coles, & Amir, 1998) ou o *Clark-Beck Obsessive-Compulsive Inventory* (Clark, Martin, Beck, Swinson, & Steer, 2005).

Em 2008, Rosa e Baptista, compuseram um instrumento para avaliar sintomas obsessivo-compulsivos, com o objetivo de obter um inventário suficientemente amplo e abrangente, a fim de avaliar vários sintomas da POC, uma vez que os distintos inventários disponíveis se dedicavam apenas a alguns dos sintomas característicos da POC. Os itens derivaram de medidas existentes, concretamente do Inventário de Padua (Burns et al., 1996), do *Maudsley Obsessive Compulsive Inventory* (Hodgson & Rachman, 1977), do *Obsessive-Compulsive Inventory* (Foa et al., 1998), do *Symetry, Ordering and Arranging Questionnaire* (Radomsky & Rachman, 2004) e do *Frost Indecisiveness Scale* (Frost & Shows, 1993). Vários estudos exploratórios foram realizados usando amostras de estudantes e de sujeitos com POC e perturbação de pânico, sendo que os resultados demonstraram que o inventário apresentou boas características psicométricas e foi eficaz na discriminação entre indivíduos com POC e sem POC (Rosa & Baptista, 2008). Como tal, poderá pensar-se que a avaliação conjunta de um maior número de sintomas, não contribui para a diminuição da confiabilidade da escala.

Ainda que os instrumentos de autorrelato demonstrem, em geral, características psicométricas adequadas, estas tendem a ser obtidas com recurso à Teoria Clássica dos Testes, parecendo necessário determinar a sua qualidade através de modelos mais avançados, tal como o Modelo de Rasch (MR) (Bond & Fox, 2007). O modelo de Rasch cumpre os requisitos de medição invariante: a medida

das pessoas não depende dos itens empregues, as pessoas mais aptas terão maior probabilidade de superar qualquer item que as menos aptas, os valores dos itens são independentes das pessoas medidas, qualquer pessoa deverá ter maior probabilidade de superar um item fácil que um item difícil e, finalmente, os itens e as pessoas localizam-se numa escala conjunta (Engelhard, 2013). Por se tratar de um modelo de medição invariante, o MR apresenta evidentes vantagens em relação à Teoria Clássica dos Testes tais como a especificidade do erro de medida e a escala de intervalo, o que justifica a realização de análises paramétricas (Prieto & Delgado, 2003).

O objetivo desta pesquisa é, então, analisar o inventário de sintomas obsessivo-compulsivos (ISOC) utilizando o modelo de medição conjunta dos itens e pessoas, especificamente, o *Rasch Rating Scale Model* (RRSM), uma versão do Modelo de Rasch para itens politómicos. O RRSM permite analisar a qualidade psicométrica das categorias de resposta em escalas de tipo *Likert*.

MÉTODO

Participantes

Participaram da pesquisa 220 indivíduos, com idades entre 18 e 62 anos. Destes, 87 são indivíduos com diagnóstico de Perturbação Obsessivo-Compulsiva, enquanto 133 são indivíduos da população geral, que não estão em qualquer processo de psicoterapia. A idade média para os indivíduos do grupo com POC é 31.45 ($DP = 10.38$) e para os indivíduos do grupo sem POC é 30.80 ($DP = 9.89$). No que diz respeito aos anos de estudo, o grupo com POC apresenta uma média de 13.16 ($DP = 3.24$) e o grupo sem POC apresenta uma média de 13.26 ($DP = 1.20$). Não se verificaram diferenças de médias estatisticamente significativas entre grupos na idade ($t(218) = -.230$; $p = .818$) e escolaridade ($t(213) = -.275$; $p = .784$).

Instrumentos

Symptom Checklist Revised (SCL-90-R). Desenvolvido por Derogatis (1977) para avaliar diferentes indicadores de psicopatologia, é composto por 90 itens divididos em nove escalas: somatização, transtorno obsessivo-compulsivo, sensi-

bilidade interpessoal, depressão, ansiedade, hostilidade, ansiedade fóbica, ideação paranoide e psicoticismo. As categorias da resposta variam de 0 (nunca) a 4 (sempre). A validação para a população portuguesa foi conduzida por Baptista (1993), que manteve a estrutura interna original, obtendo valores de consistência interna (alfa de *Cronbach*) superiores a .74 e de teste-reteste (correlação de Pearson) maior que .58. É um instrumento clínico amplamente utilizado em Portugal.

Instrumento de Avaliação de Sintomas Obsessivo-Compulsivos. Apresentado por Rosa e Baptista em 2008 com o objetivo de avaliar sintomas obsessivo-compulsivos. A sua versão atual é composta de 105 itens, retirados de diferentes instrumentos de avaliação, respondidos em formato-tipo *Likert* de 5 pontos, de 0 (nenhum) a 4 (extremamente). Avalia os sintomas obsessivos compulsivos em nove facetas características da POC; Lavagem, Verificação, Impulsos acerca de causar dano a si ou terceiros, Lentidão obsessiva, Simetria, Armazenamento, Neutralização, Dúvida obsessiva e Pensamentos sobre causar dano a si ou a terceiros.

Uma análise inicial encontrou uma alta correlação entre as facetas e o elevado valor de consistência interna para o inventário total, levando, apesar da diversidade dos conteúdos de suas facetas, a hipotetizar que o ISOC apresenta um constructo unidimensional, permitindo a sua análise usando o Modelo de Rasch.

Procedimento

Para obtenção de dados dos sujeitos pertencentes ao grupo com POC procedeu-se ao contacto com psiquiatras e psicólogos que, após autorização dos pacientes aos quais atribuíram diagnóstico de POC, lhes forneceram o protocolo de avaliação para preenchimento. O grupo de participantes sem POC foi contactado em meio académico e/ou laboral. O preenchimento do protocolo demorou entre 15 e 40 minutos. A todos os participantes foi garantido o anonimato e a confidencialidade, tendo sido informados sobre a possibilidade de desistir do preenchimento do questionário. Os dados recolhidos foram introduzidos na base de dados *Excel*, posteriormente exportados para a base de dados do *IBM SPSS Statistics 19* para comparação de médias e correlações com SCL-90-R. O tratamento estatístico do inventário com o Modelo de Rasch foi efetuado pelo *Winsteps 3.71.0* (Linacre, 2011).

O modelo de Rasch. O RRSM para escalas de qualificação (Andrich, 1988; Wright, & Mok, 2004), por ser um modelo de tipo Rasch, cumpre os requisitos de medição invariante (Engelhard, 2013). A escala, cuja unidade é o *logit*, pode variar entre menos e mais infinito, embora na maioria dos casos os valores se situem entre ± 5 *logit*. A escala de zero *logit* é convencionalmente o valor médio de localização dos itens. A unidimensionalidade é um pressuposto central do modelo. Este princípio

estabelece que a probabilidade de resposta de uma pessoa a um item é função da sua diferença numa dimensão básica (Prieto & Velasco, 2003).

Para fornecer evidência do ajuste dos dados ao modelo, duas estatísticas principais são usadas: *Outfit* e *Infit* (Linacre, 2002). Convencionalmente é considerado que a oscilação entre 0.5 e 1.5 *logit* mostra um ajuste adequado ao modelo. Valores entre 1.5 e 2.0 revelam um desajuste moderado, com poucas consequências para a validade das medidas, índices inferiores a 0.5 podem sugerir *overfit* ou redundância (menor variabilidade do que o esperado ao acaso) (López-Pina, Olivares, & Sánchez-García, 2008) e os valores mais elevados a 2.0 mostram um desajuste severo (Linacre, 2011).

Do ponto de vista do Modelo de Rasch, a confiabilidade é avaliada a nível das amostras das pessoas, designada por *Person Separation Reliability* (PSR) e dos itens - *Item Separation Reliability* (ISR). Estas estatísticas, que podem variar entre 0 e 1, têm um significado semelhante ao coeficiente de fiabilidade clássico (razão entre a verdadeira variância e a variância observada). Para uma medição correta é aconselhável que o seu valor seja superior a .70 (Cadavid, Delgado, & Prieto, 2007).

Uma condição essencial para a validade dos resultados dos testes entre grupos diferentes (masculino/feminino, clínico/não clínico) é a ausência de Funcionamento Diferencial do Item (DIF). Um item mostra DIF quando os indivíduos de grupos diferentes com o mesmo valor na variável de medida não apresentam a mesma probabilidade de uma resposta determinada. A existência de DIF no Modelo Rasch é testada através do cálculo da diferença entre as estimativas do parâmetro da dificuldade ou severidade do item em cada grupo, controlando as possíveis diferenças entre os grupos na variável latente. Wright e Douglas (1976) constataram que as diferenças menores de 0.50 *logit* têm pouco impacto na validade das medidas. Razão pela qual é geralmente considerado que o DIF é substancial se a diferença absoluta for maior do que 0.50 *logit* e estatisticamente significativa. Para testar a significância é utilizado um valor de *t de Welch* com correção de Bonferroni (Benjamini & Hochberg, 1995). Se qualquer um dos valores *t* obtiver um valor associado de $p < .05/\text{número de contrastes}$, então a hipótese de não existir funcionamento diferencial do item DIF (correção de Bonferroni) é rejeitada, sendo recomendada a sua remoção (Prieto & Delgado, 2010).

Se a resposta ao item pode ser quantificada com mais de dois valores (tal como nas escalas de tipo *Likert*), deve ser empregue qualquer um dos Modelos de Rasch para itens politómicos, entre os quais o *Rasch Rating Scale Model* (Andrich, 1988; Wright & Mok, 2004).

Linacre (2002) publicou um conjunto de recomendações para obtenção de categorias adequadas, em que o primeiro requisito é a correlação item-medida com resultados positivos e moderadamente altos. A primeira recomendação estabelece

que devem existir pelo menos 10 observações em cada categoria de resposta. A segunda recomendação diz respeito à distribuição regular de observações. A terceira linha de orientação refere-se ao aumento contínuo da média das medidas daqueles que escolhem as categorias. A quarta recomendação indica que os valores de *Outfit* de cada categoria devem ser inferiores a 2.0 (as categorias não devem apresentar um desajuste grave). Finalmente, a recomendação cinco indica que os passos entre as categorias devem ser ordenados de forma crescente. Para uma otimização das categorias de resposta recomenda-se a combinação de categorias de resposta adjacentes (Linacre, 2002).

RESULTADOS

O estudo das qualidades do ISOC iniciou-se com a análise das categorias de resposta, seguido das estatísticas de ajuste e análise do funcionamento diferencial, tendo passado por três fases - com 105 (Versão 1), 83 (Versão 2) e 78 itens (Versão 3).

Versão 1

Como Linacre (2002) sugere, procedeu-se à análise das categorias de resposta do ISOC. Primeiramente observou-se que as correlações são positivas e moderadamente elevadas (variando entre $r = .31$ e $r = .62$), sugerindo que os itens têm em comum a mesma variável latente.

A primeira recomendação de Linacre foi amplamente cumprida, uma vez que superou o número recomendado de observações por categoria (a frequência mais baixa é de 419). No que diz respeito à segunda recomendação, observa-se uma distribuição adequada (unimodal com maior frequência na categoria 1). A terceira recomendação também é cumprida, uma vez que há um aumento da média na escala das pessoas que escolhem cada categoria. Além disso, nenhuma das categorias tem um desajuste grave, apresentando todas valores de *Outfit* inferiores a 2.0 (recomendação quatro). Finalmente, é de notar que os passos entre categorias sucessivas são ordenados continuamente (recomendação cinco). Como indicado anteriormente, todas as categorias são funcionais, sendo cada uma a escolha mais provável em algum setor da variável latente.

Dado que as categorias de resposta funcionam corretamente a partir de um ponto de vista métrico, é conveniente proceder à análise das pontuações dos itens e das pessoas: ajuste ao modelo, confiabilidade e distribuição.

Qualidades métricas do ISOC105

Em relação ao ajuste dos itens ao Modelo Rasch, observa-se uma média dos valores de *Infit* de 1.06, aproximando-se do ajuste perfeito (1.00). A variabilidade é baixa ($DP = 0.19$), variando de .67 a 1.67. Para o *Outfit* são observados valores médios de 1.02 ($DP = 0.22$) e uma variação entre 0.66 e 1.97. Apenas 2.86% dos itens tem um desajuste moderado (*Infit* e/ou *Outfit* entre 1.5 e 2.0) e nenhum se revela severamente desajustado (*Infit* e/ou *Outfit* superior a 2.0), sugerindo um bom ajuste dos itens do modelo.

As médias do ajuste estatístico de pessoas estão muito perto do ajuste perfeito: 1.04 ($DP = 0.40$) para *Infit* e 1.02 ($DP = 0.44$) para o *Outfit*. As percentagens de pessoas com respostas desajustadas são pequenas: 5.91% das pessoas apresentam um desajuste moderado (> 1.5 e < 2.0) e 6.36% apresentam um alto desajuste (> 2.0).

Embora a variabilidade dos parâmetros de localização dos itens não seja elevada ($M = 0.00$; $DP = 0.057$; variação entre -1.48 e 1.49), é possível medir uma gama suficiente da variável latente, uma vez que os itens são politômicos.

Para os parâmetros das pessoas observamos valores entre -5.81 e 0.45 ($M = -1.48$; $DP = 0.82$), sugerindo que, mesmo surgindo uma variabilidade moderadamente alta, o nível das pessoas integradas na amostra total é baixa, algo esperado pela maior frequência de pessoas no grupo sem POC.

A confiabilidade das medidas dos itens e das pessoas é apropriado, excedendo o valor recomendado para uma confiabilidade minimamente aceitável ($.70$) = .97 ISR (item separation reliability) = .97; PSR (persona separation reliability) = .96 e α de Cronbach = .97.

Análise do Funcionamento Diferencial dos Itens (DIF) ISOC105

A ausência de DIF é uma condição necessária para a validade das pontuações, sendo aconselhável remover os itens que apresentem funcionamento diferencial, a fim de obter um teste com validade geral para diferentes populações de sujeitos. A análise do DIF foi realizada em relação à psicopatologia (comparação dos grupos com POC e sem POC) e sexo dos participantes do estudo.

Análise de DIF entre os grupos (com e sem POC) dos 105 itens do ISOC

Observou-se que existem 19 itens com funcionamento diferencial, uma vez que a diferença entre os parâmetros de severidade é superior a meio *logit* (.5 *logits*) e é estatisticamente significativa após a aplicação de correção de Bonferroni. Desses itens, 12 foram mais difíceis no grupo normal (e.g., item 4: “Quando ouço falar de um desastre eu acho que, de alguma forma, a culpa é minha”, enquanto itens

como 1: “Eu sinto as minhas mãos sujas quando toco em dinheiro”, apresentam menor severidade no grupo normal).

Análise de DIF entre os sexos aos 105 itens do ISOC

Os resultados da análise de funcionamento diferencial entre os itens associados com o sexo revelaram nove itens com DIF. Alguns dos itens já apresentavam DIF associado à patologia. As mulheres revelam maior severidade em itens como o 23 (“Às vezes eu tenho a tentação de rasgar minhas roupas em público”).

Versão 2

Análise de DIF entre os grupos (com e sem POC) dos 83 itens do ISOC

Conseqüentemente, foi decidido retirar os 22 itens com DIF. Porque os itens excluídos pertenciam a quase todas as facetas do inventário, pode afirmar-se que o DIF não se encontra especificamente relacionado com qualquer uma dessas facetas e que permanecem no inventário suficientes itens das mesmas. Foram eliminados itens das facetas: contaminação (“Tenho dificuldade em tocar as coisas no lixo ou sujidade”), verificação (“Passo muito tempo durante o dia a fazer verificações repetidamente”), impulsos sobre causar dano ao próprio ou terceiros (“Às vezes eu tenho o desejo de ferir crianças ou animais indefesos”), ordenação e simetria (“Sinto-me tranquilo e calmo quando os objetos à minha volta se encontram organizados e dispostos corretamente”), armazenamento (“Tenho muitas coisas que nunca uso”), dúvida obsessiva (“Habitualmente preocupo-me em fazer a escolha certa”) e pensamentos sobre causar dano ao próprio ou terceiros (“Quando eu vejo armas fico ativado e tenho pensamentos violentos”). Nenhum dos itens pertencentes a lentidão obsessiva e neutralização mostraram funcionamento diferencial. Os itens 4, 23, 28, 43, 83 e 102 apresentaram DIF entre grupo e sexo. Os itens 74, 91 e 95 apresentaram DIF entre sexos, enquanto os restantes apresentaram DIF entre os grupos com e sem POC.

Smith (2004) recomenda a reanálise dos dados após a remoção de itens com DIF. Eliminando itens que apresentam DIF em sucessivas iterações do processo de deteção, procedimento denominado de purificação, garante-se a obtenção de uma ferramenta de diagnóstico precisa e com validade. Logo, após a decisão de eliminar 22 itens da escala original, foi conduzida uma análise de DIF no ISOC reduzido (83 itens).

Desta reanálise, cinco itens revelaram funcionamento diferencial entre o grupo com POC e sem POC (correção de Bonferroni: $p \leq .0006$). Os itens excluídos correspondem às

seguintes áreas de conteúdo: o item 81 à poluição, 63 à lentidão obsessiva, 59 à neutralização, 68 à dúvida obsessiva e 52 a pensamentos sobre causar dano ao próprio ou terceiros.

Análise de DIF entre os sexos aos 83 itens do ISOC

Neste caso, nenhum dos itens demonstrou DIF associado ao sexo. Assim, decidiu-se remover os cinco itens que apresentaram DIF entre grupos e reanalisar os dados, de acordo com os requisitos de Smith (2004).

Versão 3

DIF entre grupos e DIF entre sexos

Os resultados da análise de DIF na versão de 78 itens demonstram a ausência de funcionamento diferencial associado à patologia e ao sexo. Como tal, procedeu-se a nova análise da funcionalidade das categorias de resposta, bem como ao ajustamento do modelo aos itens e pessoas e fidelidade e, ainda, à distribuição dos parâmetros dos itens e sujeitos.

Características das categorias de resposta do ISOC78

Foram encontradas correlações bisseriais positivas e moderadamente elevadas (entre 40 e 62). Verificam-se mais de 10 observações em cada categoria de resposta, com a percentagem mais elevada na primeira categoria e uma distribuição adequada da eleição das categorias. Verificou-se um aumento da média das pessoas que selecionam cada categoria, uma ausência de categorias desajustadas e um aumento crescente dos passos de calibração, podendo assim verificar-se a eficácia da funcionalidade das categorias (Tabela 1).

Estatísticas sumárias de medidas pessoa e item

Os resultados dos indicadores de ajuste dos itens e pessoas para o modelo revelaram-se aceitáveis, tal como se pode observar na Tabela 2. Na estatística *Infit* foi encontrada uma média de 1.04 ($DP = 0.16$) e em relação ao *Outfit* o indicador mais sensível às respostas inesperadas revelou-se muito perto do ajuste perfeito, com um valor médio de 1.03 ($DP = 0.20$). Ainda que a variabilidade não seja alta é possível medir uma gama suficiente da variável latente, uma vez que os itens são politómicos. Os valores das medidas entre as categorias variam de -1.14 e 1.22 e

a amplitude dos indicadores para medir as pessoas oscila entre -2.08 e 2.15 ($D_i + F_k$), o que garante uma amostragem aceitável da variável.

Tabela 1
Efetividade das Categorias de Resposta do ISOC-78

Passo	Linha orientadora	Categoria	ISOC	
Pre.	Escala orientada com a variável latente	---	Sim (.41 a .63)	
1	Pelo menos 10 observações em cada categoria	Categoria 1	8208(48%)	
		Categoria 2	4989 (29%)	
		Categoria 3	2573 (15%)	
		Categoria 4	1107 (6%)	
		Categoria 5	283 (2%)	
2	Distribuição regular das observações	---	Sim	
3	Aumento contínuo das médias das medidas das pessoas que selecionam as categorias	Categoria 1	-2.00	
		Categoria 2	-1.24	
		Categoria 3	-.72	
		Categoria 4	-.43	
		Categoria 5	-.33	
			<i>Infit</i>	<i>Outfit</i>
4	<i>Outfit</i> menor que 2.0	Categoria 1	.94	.96
		Categoria 2	.97	.84
		Categoria 3	.93	.93
		Categoria 4	1.13	1.25
		Categoria 5	1.39	1.76
5	Aumento crescente dos passos	Categoria 1	---	
		Categoria 2	-1.14	
		Categoria 3	-.38	
		Categoria 4	.29	
		Categoria 5	1.22	

As médias das estatísticas de ajuste das pessoas estão muito próximas da unidade (ajuste perfeito) tanto para o *Infit* ($M = 1.03$ e $DP = 0.44$) como para o *Outfit* ($M = 1.03$ e $DP = 0.47$). Apenas 19 pessoas (8.64%) apresentam um desajuste moderado (> 1.5 e < 2) e 17 (7.73%) um alto desajuste (> 2). Logo, as respostas da maioria dos participantes não foram diferentes daquelas previstas pelo modelo. Para os parâmetros das pessoas observaram-se valores entre -3.98 e 0.63 ($M = -1.46$, $DP = 0.85$), sugerindo que, apesar de uma variabilidade moderadamente alta, o nível das pessoas integradas na amostra total parece baixo, provavelmente devido à menor frequência de pessoas com POC.

A precisão ou confiabilidade das medidas dos itens e das pessoas é adequada, excedendo o valor recomendado para uma confiabilidade minimamente aceitável (.70): $ISR = .96$; $PSR = .96$ e α de *Cronbach* = .97.

Tabela 2
Resumo das Medidas de Pessoas e Itens para o ISOC 78

Estatísticas	Min.	Max	M	DP	Valor
<i>Infit</i> dos itens	.69	1.42	1.04	.16	
<i>Outfit</i> dos itens	.67	1.59	1.03	.20	
<i>Infit</i> das pessoas	.33	2.69	1.03	.44	
<i>Outfit</i> das pessoas	.35	2.77	1.03	.47	
% de itens com desajuste moderado (1)					3.85%
% de itens com desajuste elevado (2)					0%
% de pessoas com desajuste moderado (1)					8.64%
% de pessoas com desajuste elevado (2)					7.73%
Parâmetros dos itens	-.94	.93	.00	.45	
Parâmetros das pessoas	-3.98	.63	-1.46	.85	
Confiabilidade dos itens					.96
Confiabilidade das pessoas					.96
Alfa <i>Cronbach</i>					.97
(1)	<i>Infit</i> e/ou <i>Outfit</i> <.5; >1.5<2.0				
(2)	<i>Infit</i> e/ou <i>Outfit</i> > 2.0				

Escalamento conjunto de pessoas e itens

Uma vantagem do Modelo de Rasch é a possibilidade de uma medição conjunta de itens e pessoas, já que os seus parâmetros estão na mesma escala. Na Figura 1 pode observar-se graficamente o escalamento conjunto. As pessoas estão distribuídas entre 0.63 e -3.98 *logit*, enquanto os itens são entre 0.93 e -0.94 *logit*. Tal como supramencionado, uma vez que a maioria das pessoas tem pontuações mais baixas do que 0 (a severidade média dos itens), o nível dos participantes no atributo medido é baixo. Tendo em conta que estes são indicadores de psicopatologia, este resultado é consistente com as expectativas.

Logit	Pessoa	Items								
1		32	94							
	.	104	54	67	72	88				
		38	76	79	87					
	#	100	48	85						
		22	42	44	62	75	93			
	#	49	5	69	71	82	84			
		10	39	41	50	57	58	61	77	97
0	##	16	18	19	26	30	47	56	78	
	.#	101	11	36	55	70	9	92	96	
	##	12	37	6	66	89				
	#####	20	34	46	60	73	8	98		
	####	15	27	31	33	53				
	.#####	13	2	35						
	#####	24								
-1	####	7								
	.#####									
	#####									
	.#####									
	#####									
	.#####									
	.#####									
-2	.#####									
	####									
	####									
	###									
	###									
	.##									
	.									
-3	###									
	.#									
	.									
	.									
	.									
-4	.#									

Nota: Cada . representa uma pessoa e cada # representa duas pessoas

Figura 1. Escalamento conjunto de pessoas e itens.

Os valores médios (logit) das diferentes facetas

Devido ao ajustamento estatístico, verificou-se que os resultados observados pouco diferem daqueles previstos pelo modelo. Uma vez que as respostas esperadas envolvem a unidimensionalidade (a resposta é uma função da diferença de uma dimensão entre a pessoa e o item), é razoável concluir que o ISOC apresenta um ajustamento aceitável ao pressuposto de unidimensionalidade. Tal não é incompa-

tível com o facto de as distintas facetas de conteúdo apresentarem diferentes graus de severidade na variável medida.

Foi encontrada a seguinte ordem de severidade das facetas: impulsos de causar dano ao próprio ou terceiros (NItens = 2; $M = 0.78$; $DP = 0.07$), indicando a maior gravidade da POC, neutralização (NItens = 6; $M = 0.39$; $DP = 0.41$), lentidão (NItens = 4; $M = 0.34$; $DP = 0.43$), dúvida (NItens = 11; $M = 0.32$; $DP = 0.41$), contaminação (NItens = 12; $M = 0.20$; $DP = 0.40$), verificação (NItens = 9; $M = 0.16$; $DP = 0.35$), armazenamento (NItens = 12; $M = 0.12$; $DP = 0.44$), ordenação (NItens = 16; $M = 0.11$; $DP = 0.36$) e pensamento de causar dano ao próprio ou terceiros (NItens = 6; $M = 0.06$; $DP = 0.46$), revelando o menor nível da perturbação ($M = 0.06$; $DP = 0.46$). Existem diferenças significativas entre as médias ($F = 3.84$, $df = 8$, $p = .001$), porém, dado o pequeno número de itens em algumas facetas, estudos de *post-hoc* não foram realizados.

Estatísticas descritivas e diferenças médias (logit) entre os grupos

A análise das estatísticas descritivas do grupo clínico e o grupo normal indicou uma média de -0.89 ($DP = 0.08$) no grupo clínico e uma média de -1.87 ($DP = 0.07$) no grupo normal. A Estatística t de Welch para testar a diferença entre as médias foi estatisticamente significativa ($\alpha = .05$) e de grande tamanho ($Welch t(202) = 9.28$, $p < .001$). A distância foi de $.98$ *logits* (d de Cohen = 1.24 um valor muito elevado de tamanho do efeito).

Estatísticas descritivas e diferenças médias (logit) entre homens e mulheres

Embora os escores médios de homens ($M = -1.32$; $DP = 0.12$) e mulheres ($M = -1.58$; $DP = 0.07$) tenham diferido a nível descritivo, as diferenças não foram estatisticamente significativas ($Welch t(141) = 1.96$; $p = .052$).

Correlação do ISOC com as subescalas da SCL-90-R

A pontuação total do ISOC foi correlacionada com as subescalas da SCL-90-R, utilizando a correlação de Pearson. Os valores encontrados revelaram-se estatisticamente significativos ($p \leq .01$) variando de moderados a altos (entre $.49$ e $.59$), sobressaindo o valor de $.71$ ($p < .001$) com a subescala de obsessões e compulsões.

DISCUSSÃO

O objetivo deste trabalho foi analisar as características psicométricas do ISOC, utilizando o Modelo de Rasch para escalas de classificação. Após eliminação de 27 itens com DIF significativo, foi efetuada uma nova análise das categorias de resposta e das qualidades psicométricas da versão de 78 itens. Da análise das categorias de resposta encontraram-se resultados que indicam a adequação das cinco possibilidades de resposta. As estatísticas de ajuste dos itens apresentaram valores de *Infit* e *Outfit* em torno da unidade, mostrando um comportamento adequado (Linacre, 2002). Em relação à confiabilidade, os resultados proporcionam excelentes índices para os itens e as pessoas, bem como um α de *Cronbach* muito elevado, indicando uma precisão adequada do instrumento (Cadavid, Delgado, & Prieto, 2007). Quanto às correlações ponto bisserial, os valores positivos indicam que os itens são orientados com a variável latente e, deste modo, compartilham o mesmo constructo (Mok, Cheng, Moore, & Kennedy, 2004).

Não se observou uma grande amplitude das pontuações das pessoas, indicando pouca amplitude de sintomas obsessivo-compulsivos na amostra total. A média da gravidade dos itens excede o nível apresentado pelos sujeitos, o que mostra a sua utilidade na avaliação de uma variável de natureza clínica como a POC.

Os dados do ISOC estão em conformidade com o Modelo de Rasch, uma vez que cumprem os critérios de qualidade propostos pelo Linacre (2002), sugerindo um instrumento de avaliação dos sintomas obsessivos compulsivos preciso e com melhores propriedades psicométricas do que as de outras versões.

A correlação com as dimensões da SCL-90-R foi concebida para fornecer um indicador externo de validade, esperando-se relações mais fortes entre o ISOC e a dimensão das obsessões e compulsões, comparativamente com os valores de correlação com as outras dimensões da psicopatologia, o que se verificou.

A análise do impacto associado ao sexo não revelou resultados significativos. A literatura consultada no que respeita às diferenças entre os sexos não é unânime, referindo igual prevalência entre os sexos ou um valor ligeiramente superior nas mulheres (APA, 2013; Torresan et al., 2009). Quanto ao impacto associado com o grupo, foram encontradas diferenças estatisticamente significativas entre os grupos (normal e clínico). É de destacar o grande tamanho do efeito, mais de um desvio padrão, associado a esta diferença. Tal resultado vai na direção esperada, uma vez que foi proposto proporcionar uma medida eficaz para discriminar entre sujeitos com e sem POC, a qual possa ser utilizada na prática clínica e na investigação para avaliar a gravidade dos sintomas e/ou para medir o resultado de tratamentos (Blom, Larsson, Serlachius, & Ingvar, 2010).

Embora os resultados indiquem que o ISOC pode ser considerado unidimensional de um ponto de vista estatístico, algo que acarreta vantagens psicométricas, o conteúdo dos itens permite-nos levar em consideração as diferentes facetas da

POC (por exemplo, lavagem, verificação ou dúvida) para uma análise individualizada das respostas dos pacientes. Comparativamente com a versão anterior (2008) o ISOC revela-se um instrumento mais robusto, já que as fortes correlações entre as distintas dimensões sugeriam a existência de dependência entre as mesmas.

REFERÊNCIAS

- American Psychiatric Association (APA) (2013). *Diagnostic and statistical manual of mental disorders DSM 5*. Arlington, V.A.: American Psychiatric Publishing.
- Andrich, D. (1988). *Rasch models for measurement*. London: Sage Publications.
- Baptista, A. (1993). *A gênese da perturbação de pânico: A importância dos factores familiares e ambientais durante a infância e adolescência* (Tese de doutoramento não publicada), Universidade do Porto, Porto.
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, 57(1), 289-300.
- Blom, E. H., Larsson, J.-O., Serlachius, E., & Ingvar, M. (2010). The differentiation between depressive and anxious adolescent females and controls by behavioural self-rating scales. *Journal of Affective Disorders*, 122(3), 232-240. doi: 10.1016/j.jad.2009.07.006
- Bond, T. G., & Fox, C. M. (2007). *Applying the Rasch model: Fundamental measurement in the human sciences* (2nd ed.). Mahwah: Lawrence Erlbaum Associates, Publishers
- Burns, G. L., Keortge, S., Formea, G., & Sternberger, L. (1996). Revision of the Padua Inventory of obsessive-compulsive disorder symptoms: Distinctions between worry, obsessions, and compulsions. *Behaviour Research and Therapy*, 34(2), 163-173.
- Cadavid, N., Delgado, A. R., & Prieto, G. (2007). Construcción de una escala de depresión con el modelo de Rasch. *Psicothema*, 19(3), 515-521. Consultado em <http://www.psicothema.com/pdf/3393.pdf>
- Clark, D. A., Martin, A. M., Beck, A. T., Swinson, R. P., & Steer, R. A. (2005). Screening for obsessive and compulsive symptoms: Validation of the Clark-Beck Obsessive-Compulsive Inventory. *Psychological Assessment*, 17(2), 132-143. doi: 10.1037/1040-3590.17.2.132
- Cooper, J. (1970). The Leyton Obsessional Inventory. *Psychological Medicine*, 1(1), 48-64. doi: 10.1017/S0033291700040010
- Derogatis, L. R. (1977). *Symptom Checklist-90: Administration, scoring and procedures manual for the revised version*. Baltimore, MD: John Hopkins University School of Medicine, Clinical Psychometrics Research Unit.
- Engelhard, G. (2013). *Invariant measurement: Using Rasch models in the social, behavioral and health sciences*. New York, NY: Routledge.
- Foa, E. B., Kozak, M. J., Salkovskis, P. M., Coles, M. E., & Amir, N. (1998). The validation of a new obsessive compulsive disorder scale: The Obsessive Compulsive Inventory (OCI). *Psychological Assessment*, 10(3), 206-214. doi: 10.1037/1040-3590.10.3.206
- Franklin, M. E., & Foa, E. B. (2011). Treatment of obsessive-compulsive disorder. *Annual Review of Clinical Psychology*, 7, 229-243. doi: 10.1146/annurev-clinpsy-032210-104533
- Frost, R. O., & Shows, D. L. (1993). The nature and measurement of compulsive indecisiveness. *Behaviour Research and Therapy*, 31(7), 683-692. doi: 10.1016/0005-7967(93)90121-A

- Fullana, M. A., Vilagut, G., Rojas-Farreras, S., Mataix-Cols, D., Graaf, R., Demyttenaere, K., ... Alonso, J. (2010). Obsessive-compulsive symptom dimensions in the general population: Results from an epidemiological study in six European countries. *Journal of Affective Disorders, 124*(3), 291-299. doi: 10.1016/j.jad.2009.11.020
- Hodgson, R. J., & Rachman, S. (1977). Obsessional-compulsive complaints. *Behaviour Research and Therapy, 15*(5), 389-395. doi: 10.1016/0005-7967(77)90042-0
- Linacre, J. M. (2002). Optimizing rating scale category effectiveness. *Journal of Applied Measurement, 3*(1), 85-106. Consultado em <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.424.2811&rep=rep1&type=pdf>
- Linacre, J. M. (2011). *A user's guide do Winstep Ministep: Rasch-model computer program*. winsteps.com
- López-Pina, J. A., Olivares, J., & Sánchez-García, R. (2008). Rasch modeling of the Spanish self-report version of the Liebowitz Social Anxiety Scale for Children and Adolescents (LSAS-CA-SR). *International Journal of Clinical and Health Psychology, 8*(1), 233-245. Consultado em <http://www.redalyc.org/articulo.oa?id=33780116>
- McKay, D., Abramowitz, J. S., Calamari, J. E., Kyrios, M., Radosky, A., Sookman, D., ... Wilhelm, S. (2011). A critical evaluation of obsessive-compulsive disorder subtypes: Symptoms versus mechanisms. *Clinical Psychology Review, 24*(3), 283-313. doi: 10.1016/j.cpr.2004.04.003
- Mok, M., Cheng, C., Moore, P., & Kennedy, K. (2004). *The development of measurement scales on self-learning of secondary students*. Paper presented to the Australian Association for Research in Education, University of Melbourne, Australia. Consultado em <http://www.aare.edu.au/data/publications/2004/mok04933.pdf>
- Prieto, G., & Delgado, A. R. (2003). Análisis de un test mediante el modelo de Rasch. *Psicothema, 15*(1), 94-100. Consultado em <http://www.psicothema.com/pdf/1029.pdf>
- Prieto, G., & Delgado, A. R. (2010). Fiabilidad y validez. *Papeles del Psicólogo, 31*(1), 67-74. Consultado em <http://www.papelesdelpsicologo.es/pdf/1797.pdf>
- Prieto, G., & Velasco, A. (2003). Uso del modelo de Rasch para poner en la misma escala las puntuaciones de distintos tests. *Actualidades en Psicología, 19*(106), 5-23. doi: 10.15517/ap.v19i106.43
- Radosky, A. S., & Rachman, S. (2004). Symmetry, ordering and arranging compulsive behaviour. *Behaviour Research and Therapy, 42*(8), 893-913. doi: 10.1016/j.brat.2003.07.001
- Rosa, J. B., & Baptista, A. (2008). Análise das características psicométricas do inventário obsessivo-compulsivo. *Actas da Conferência Avaliação Psicológica: Formas e Contextos*. Braga: Psiquilíbrios.
- Smith, R. M. (2004). Detecting item bias with the Rasch model. In E. V. Smith Jr & R. M. Smith (Eds.). *Introduction to Rasch Measurement* (pp. 391-418). Maple Grove, MN: JAM Press.
- Steketee, G., & Freund, B. (1993). Compulsive Activity Checklist (CAC): Further psychometric analyses and revision. *Behavioural Psychotherapy, 21*(1), 13-25. doi: 10.1017/S0141347300017766
- Thordarson, D. S., Radosky, A. S., Rachman, S., Shafran, R., Sawchuk, C. N., & Ralph Hakstian, A. (2004). The Vancouver Obsessional Compulsive Inventory (VOCI). *Behaviour Research and Therapy, 42*(11), 1289-1314. doi: 10.1016/j.brat.2003.08.007
- Torresan, R. C., Ramos-Cerqueira, A. T. A., Mathis, M. A., Diniz, J. B., Ferrão, Y. A., Miguel, E. C., & Torres, A. R. (2009). Sex differences in the phenotypic expression of obsessive-compulsive disorder: An exploratory study from Brazil. *Comprehensive Psychiatry, 50*(1), 63-69. doi: 10.1016/j.comppsy.2008.05.005
- Wright, B. D., & Douglas, G. A. (1976). Rasch item analysis by hand. *Research Memorandum, 21*, Statistical Laboratory, Department of Education, University of Chicago.
- Wright, B. D., & Mok, M. M. (2004). An overview of the family of Rasch measurement models. In E. V. Smith Jr & R. M. Smith, *Introduction to Rasch Measurement* (pp. 1-24). Maple Grove, MN: JAM Press.