

# Inventário de Depressão de Beck II: Análises pela Teoria do Traço Latente

Marli Appel da Silva<sup>1</sup>

Pontifícia Universidade Católica do Rio Grande do Sul – PUC/RS, Porto Alegre-RS, Brasil

Guilherme Welter Wendt

University of London, Londres, Reino Unido

Irani Iracema de Lima Argimon

Pontifícia Universidade Católica do Rio Grande do Sul – PUC/RS, Porto Alegre-RS, Brasil

## RESUMO

Este estudo tem o objetivo de avaliar as propriedades psicométricas do Inventário de Depressão de Beck – Segunda Versão (BDI-II), que visa avaliar a gravidade dos sintomas depressivos. A amostra contou com 2.385 participantes (13 a 85 anos) da região metropolitana de Porto Alegre, Rio Grande do Sul. Os seguintes métodos da Teoria do Traço Latente foram utilizados: Modelagem de Equação Estrutural Exploratória, Análise de Classe Latente e Análise Fatorial Confirmatória Multigrupo. Observou-se a possibilidade de que o BDI-II possa ser um instrumento unidimensional com subfatores mais restritos (modelo bifator). Encontrou-se evidências de que o instrumento apresente precisão de medida para a mensuração da gravidade dos sintomas depressivos. Também, pontos de corte foram sugeridos, verificando a possibilidade de três classes em termos de gravidade dos sintomas depressivos: mínimos, leves e moderados/graves. Dentre as limitações deste estudo, destaca-se que não foi efetuada a comparação com outras amostras com características diferenciadas.

*Palavras-chave:* sintomas depressivos; BDI-II; Teoria do Traço Latente.

## ABSTRACT – Beck Depression Inventory II: analyses by Latent Trace Theory

This study aims to evaluate psychometric properties of the Beck Depression Inventory – Second Version (BDI-II), which aims to evaluate the severity of depressive symptoms. The following methods of the Latent Trace Theory were used: Exploratory Structural Equation Modeling, Latent Class Analysis and Multi-group Confirmatory Factor Analysis. The possibility that BDI-II could be a one-dimensional instrument with more restricted subfactors (bi-factor model) was observed. There was evidence that the instrument presented precision in measuring the severity of depressive symptoms. Cutoff points were also suggested, verifying the possibility of three classes in terms of depressive symptoms severity: minimal, mild, and moderate/severe. Among the limitations of this study, it should be noted that the comparison with other samples with different characteristics was not performed.

*Keywords:* depressive symptoms; BDI-II; latent trace theory.

## RESUMEN – Inventario de Depresión de Beck II: análisis por la Teoría del Trazo Latente

Este estudio tiene el objetivo de evaluar las propiedades psicométricas del Inventario de Depresión de Beck – Segunda Versión (BDI-II), que busca evaluar la gravedad de los síntomas depresivos. La muestra contó con 2.385 participantes (13 a 85 años) de la región metropolitana de Porto Alegre, en la provincia de Rio Grande do Sul. Se utilizaron los siguientes métodos de la Teoría del Trazo Latente: Modelo de Ecuación Estructural Exploratoria, Análisis de Clase Latente y Análisis Fatorial Confirmatorio Multigrupo. Se observó la posibilidad de que el BDI-II pueda ser un instrumento unidimensional con subfactores más restringidos (modelo bifactor). Quedó evidente la precisión del instrumento utilizado para medir la gravedad de los síntomas depresivos. También, fueron sugeridos puntos de corte, verificando la posibilidad de tres distintas clases en términos de gravedad de los síntomas depresivos: mínimos, leves y moderados/graves. Entre las limitaciones de este estudio, cabe destacar, que no se efectuó la comparación con otras muestras con características diferenciadas.

*Palabras clave:* síntomas depresivos; BDI-II; teoría del trazo latente.

Os transtornos depressivos fazem alusão a uma condição com alto potencial incapacitante (Lam, 2018). Em linhas gerais, profissionais buscam avaliar se há redução de prazer, humor deprimido, falta de disposição/energia,

perturbações no sono e apetite, dentre outros sintomas, em diversas esferas da vida quando examinam indivíduos possivelmente candidatos a possuírem algum tipo de transtorno depressivo (Cowen, 2017). Tais transtornos

<sup>1</sup> Endereço para correspondência: Rua Dr. Flores, 163, cj. 1101, 90020-120, Porto Alegre, RS. Tels.: (51) 99844-2041 / 3894-2142. E-mail: mappel@uol.com.br

englobam, em determinado período de tempo ou mesmo recorrentes, especificadas condições clínicas, incluindo os transtornos: disruptivo da desregulação do humor (explosões de raiva recorrentes e graves), maior (sintomas depressivos episódicos e recorrentes), persistente (distímia), disfórico pré-menstrual (na maioria dos ciclos menstruais), induzido por substância/medicamento (após uso, abuso ou abstinência), devido a outra condição médica (consequência fisiopatológica), especificado e não especificado (não satisfazem os critérios para outros transtornos; Associação Americana de Psiquiatria, 2014). Em muitos casos, a sintomatologia depressiva pode inclusive não envolver a presença de sintomas emocionais, o que torna imperativo que os demais componentes sejam bem conhecidos e avaliados (e.g., sintomas físicos e cognitivos; Lam, 2018).

Em 2015, a prevalência de todos os transtornos depressivos (e.g., maior, persistente, induzido por substância/medicamento) foi estimada em cerca de 4% da população global e 6% da população brasileira (World Health Organization, 2017). Em relação apenas ao transtorno depressivo maior, na população brasileira, a prevalência pode ser de 8% durante um ano e de 17% ao longo da vida. De maneira interessante, nuances socioculturais podem explicar as diferenças na prevalência dos transtornos depressivos nas regiões brasileiras (Silva, Galvao, Martins, & Pereira, 2014).

Acrescenta-se que os sintomas depressivos não estão circunscritos apenas aos transtornos depressivos, podem estar presentes em uma série de condições clínicas (e.g., estresse pós-traumático, demências; Argimon, Paloski, Farina, & Irigaray, 2016); ou, ainda, podem ser o resultado de respostas específicas e momentâneas a situações adversas ou estressantes (Ely, Nunes, & Carvalho, 2014). Um estudo de metanálise indicou que a prevalência dos sintomas depressivos autorrelatados para a população brasileira pode estar na ordem de 14% (Silva et al., 2014), apontando ainda haver ocorrência mais elevada desses sintomas entre as mulheres. Todavia, existem ainda algumas dificuldades para que se consiga apreender com maior robustez o impacto da sintomatologia depressiva entre brasileiros, destacando-se o limitado número de estudos em determinadas regiões do país, bem como a falta de uniformidade entre profissionais quando do exame clínico dos transtornos e sintomas depressivos.

### **Inventário de Depressão de Beck II: evidências acerca do desempenho psicométrico**

As escalas de depressão de Beck visam avaliar a gravidade dos sintomas depressivos, sendo consideradas ferramentas importantes no diagnóstico, no prognóstico e na evolução da sintomatologia dos pacientes, bem como na uniformização da comunicação entre profissionais (Poole, Bramwell, & Murphy, 2009). A primeira versão do inventário (BDI-I) foi desenvolvida por Beck, Ward, Mendelson, Moch e Erbaugh (1961), sendo

adaptada para a população brasileira por Cunha (2001). A segunda versão (BDI-II; Beck, Steer, & Brown, 1996) baseou-se na revisão da primeira edição com a finalidade de se alinhar aos critérios para os transtornos depressivos fornecidos pelo Manual Diagnóstico e Estatístico de Transtornos Mentais – Quarta Edição (DSM-IV; Associação Americana de Psiquiatria, 1995). Essa versão foi adaptada à população brasileira por Gorenstein, Pang, Argimon e Werlang (2011).

Por meio da análise fatorial exploratória (AFE), estudos que utilizaram a segunda versão do instrumento (BDI-II) encontraram variadas estruturas fatoriais em diferentes populações (e.g. Beck et al., 1996; Gorenstein et al., 2011; Hayes et al., 2015). A AFE é utilizada para verificar se um conjunto de variáveis mensuradas (observáveis) engloba um construto ou fator (variável latente; Damásio, 2012).

Já a análise fatorial confirmatória (AFC) objetiva avaliar as relações entre variáveis observáveis e latentes (não mensuradas diretamente) com a finalidade de estabelecer se um modelo derivado de uma determinada teoria possui a possibilidade de aproximação de uma base populacional (Hair, Black, Babin, Anderson, & Tatham, 2009). Consequentemente, em relação ao BDI-II, foram testados modelos com um, dois ou três fatores (e.g. Strunk & Lane, 2017; Wu & Huang, 2014), ou mesmo hierárquicos (depressão geral representando o fator de ordem superior; e.g. Harris & D'Eon, 2008; Whisman, Judd, Whiteford, & Gelhorn, 2013) empregando métodos da AFC.

Todavia, algumas investigações que avaliaram modelos da AFC com dois ou três fatores relacionados obtiveram covariâncias entre fatores acima de 0,60 (e.g. Argimon et al., 2016 et al., 2015; Wu & Huang, 2014). Do mesmo modo, outros estudos identificaram padrões de correlações entre os fatores acima de 0,90 para os modelos hierárquicos (e.g. Harris & D'Eon, 2008; Whisman et al., 2013), o que sugeriu que alguns itens possam mensurar aspectos multidimensionais (Rodríguez, Reise, & Haviland, 2016). Com efeito, são recomendáveis valores iguais ou menores que 0,60 e significativos ( $p \leq 0,05$ ) para as covariâncias entre fatores para que haja indícios de capacidade discriminante entre fatores; e, idealmente, acima de 0,70 e abaixo de 0,90 para as correlações entre fatores em modelos hierárquicos (Hair et al., 2009).

Não obstante, as covariâncias/correlações entre fatores podem ser elevadas em instrumentos psicológicos em decorrência das dimensões psíquicas serem interconectadas; assim, os itens podem apresentar tanto o fenômeno da unidimensionalidade (fator geral preponderante) como da multidimensionalidade (fatores específicos mais restritos; Reise, 2012; Rodríguez et al., 2016). De fato, Bringmann, Lemmens, Huibers, Borsboom e Tuerlinck (2015) verificaram que todos os sintomas depressivos citados no BDI-II estão direta ou indiretamente conectados, ou seja, formam uma rede de interações.

Quando há a indicação da existência de uni e multidimensionalidade concomitantes em um instrumento, torna-se possível testar a hipótese de modelos bifatores, contemplando um fator geral ortogonal e subfatores ortogonais que explicam a variância dos itens (Reise, 2012; Rodriguez et al., 2016). Sendo assim, modelos bifatores confirmatórios foram testados para o BDI-II, nos quais os itens foram englobados em um fator geral (depressão) e também em dois subfatores não correlacionados (e.g. Brouwer, Meijer, & Zevalkink, 2013; Dere et al., 2015; Ward, 2006). De acordo com Brouwer et al. (2013), as soluções bifatoriais foram as que encontraram mais consenso na literatura em relação ao BDI-II.

Ademais, distintos pontos de corte para determinar a gravidade dos sintomas depressivos foram propostos ao BDI-II (e.g. Kjaergaard, Arfwedson Wang, Waterloo, & Jorde, 2014; Plourde, Moullec, Bacon, Suarhana, & Lavoie, 2016). Comparando a entrevista clínica estruturada baseada no DSM-III-R com o BDI-II por meio do método Característica de Operação do Receptor (*Receiver Operating Characteristic*, ROC), em uma amostra de 127 pessoas da Universidade da Pensilvânia, o estudo de Beck et al. (1996) sugeriu os seguintes parâmetros: mínimo (0-13), leve (14-19), moderado (20-28) e grave (29-63).

Na versão do BDI-II adaptada para a população brasileira (Gorenstein et al., 2011), os pontos de corte adotados por Beck et al. (1996) foram comparados com os diagnósticos obtidos com a entrevista clínica estruturada baseada no DSM-IV. Por meio da análise discriminante, que compara uma amostra de casos com o diagnóstico conhecido com novos escores para a variável preditiva, observou-se uma maior capacidade discriminativa em relação às pontuações mais baixas (0-13 pontos; discriminação correta = 81,1%) e elevadas (29-63 pontos; discriminação correta = 85,7%). Os casos classificados como sintomas depressivos leves (14-19 pontos) e moderados (20-28 pontos) obtiveram uma discriminação correta de 50% e 77,8%, respectivamente.

Em resumo, investigações anteriores revelaram diferenciadas soluções fatoriais e pontos de corte para o BDI-II, denotando a possibilidade do inventário apresentar sensibilidade às variabilidades das populações (Brouwer et al., 2013; Strunk & Lane, 2017). Aspecto que instiga a discussão a respeito da precisão (possibilidade de invariabilidade amostral) e acurácia diagnóstica do instrumento (Coulacoglou & Saklofske, 2017; Strunk & Lane, 2017).

## O presente estudo

Considerando que as síndromes depressivas figuram como incapacitantes e causas de morte tanto no Brasil como no mundo, esforços em apurar o diagnóstico e tratamento dessas problemáticas são imperativos (World Health Organization, 2017). Nesse sentido, o BDI-II pode ser uma ferramenta útil aos profissionais tendo em vista que permite o rastreamento de casos com sintomas depressivos, auxiliando a aplicação de tratamentos

necessários e a comunicação efetiva entre os trabalhadores da saúde (Minas, 2017; Wang & Gorenstein, 2013b). Além disso, estudos prévios com o BDI-II podem ser localizados na literatura em diferenciadas populações/amostras (e.g., Bhardwaj, Price, Landry, Harvey, & Hensel, no prelo; Doshi et al., 2018), o que demonstra o valor do instrumento para a prática da pesquisa, já que pode apoiar aos pesquisadores na obtenção de dados sobre os sintomas depressivos (Lam, 2018).

Porém, caso as propriedades psicométricas das medidas e os pontos de corte utilizados não sejam adequados para as populações/amostras investigadas, pode ocasionar a detecção de falsos positivos ou negativos, impactando na saúde das populações (Strunk & Lane, 2017). Desse modo, torna-se necessário ampliar as investigações em relação às propriedades psicométricas do BDI-II para populações específicas, tendo em vista que a busca de evidências sobre as propriedades psicométricas dos instrumentos psicológicos deve ser um processo contínuo (Primi, 2010).

Mediante a esse contexto, o presente estudo tem o objetivo de avaliar, com base em técnicas da Teoria do Traço Latente (TTL), as propriedades psicométricas do BDI-II (Beck et al., 1996) em amostra proveniente da região metropolitana de Porto Alegre, Rio Grande do Sul, Brasil. A TTL apresenta o pressuposto de que os itens (variáveis observáveis ou mensuradas) possuem a capacidade de medir um processo psíquico (traço latente; Pasquali, 2009). Assim, o presente estudo valeu-se do uso da Modelagem de Equação Estrutural Exploratória (*Exploratory Structural Equation Modeling*, ESEM; Asparouhov & Muthén, 2009) para avaliar as possibilidades fatoriais para o instrumento; da Análise de Classe Latente (Muthén, 2008) com a finalidade de propor pontos de corte ao BDI-II; e da Análise Fatorial Confirmatória Multigrupo (AFCMG; Damásio, 2013; Hair et al., 2009), de modo a verificar a possibilidade de invariabilidade e variância entre os grupos ou classes pertencentes à amostra analisada. A ESEM une técnicas da AFE e AFC para a análise das melhores soluções fatoriais (Asparouhov & Muthén, 2009). Já a análise de classe latente visa, de uma amostra heterogênea, classificar subgrupos não observados ou classes latentes por meio de padrões semelhantes de dados (Berlin, Williams, & Parra, 2013; Muthén, 2008). Por sua vez, a AFCMG permite verificar a possibilidade de invariabilidade e variância entre grupos (Damásio, 2013; Hair et al., 2009).

## Método

### Participantes

A amostra contou com 2.385 participantes entre 13 e 85 anos (média de idade=26,62), residentes na região metropolitana de Porto Alegre, Rio Grande do Sul, Brasil. A maior parte tinha entre 20 a 29 anos. Em termos

de educação, o ensino médio completo foi indicado pela maioria dos participantes. Além disso, o grupo não clínico (sem diagnósticos psiquiátricos) foi preponderante na amostra. Em relação às pontuações médias no BDI-II, as maiores foram apresentadas por: mulheres, entre

indivíduos com 30 e 59 anos de idade, com ensino fundamental, pertencentes ao grupo psiquiátrico, com transtornos de ansiedade e humor. No presente estudo, a média total das respostas ( $M_r$ ) foi de 11,81 (desvio padrão das respostas [ $DP_r$ ]=17,20; Tabela 1).

Tabela 1  
Dados Sociodemográficos

Variáveis	Categorias	%	$M_r$	$DP_r$
Sexo	Masculino	45,95	8,73	8,19
	Feminino	54,05	14,24	11,61
Faixa etária	13 até 19	26,54	9,93	8,33
	de 20 a 29	36,81	10,97	10,13
	de 30 a 39	13,38	14,52	12,76
	de 40 a 49	11,45	13,86	12,19
	de 50 a 59	7,09	14,42	11,47
	60 ou mais	4,74	10,68	9,28
Escolaridade	Ensino fundamental	29,06	15,91	12,39
	Ensino médio	65,62	10,61	9,13
	Superior completo	5,32	9,57	8,84
Grupos	Psiquiátrico	6,16	21,54	12,19
	Clínico	17,23	12,64	11,46
	Não clínico	76,61	10,71	9,72
Diagnósticos	Dependência química	1,51	22,11	13,46
	Alcoolismo	2,22	17,13	8,74
	Transtorno de ansiedade	0,80	25,37	13,04
	Transtorno de humor	1,64	25,15	13,06
	Incapacidade física	10,23	14,51	10,93
	Incapacidade intelectual	7,00	9,90	11,70
	Sem diagnóstico	76,60	10,71	9,72
Total		100,00	11,81	17,20

Nota. %=percentual;  $DP_r$ =desvio padrão das respostas;  $M_r$ =média das respostas

O grupo psiquiátrico apresentou diagnósticos obtidos com base em seus prontuários médicos, a saber: dependência química, alcoolismo, transtorno de ansiedade e de humor. O grupo clínico configurou-se por pessoas com necessidades especiais (incapacidade física e intelectual), bem como o grupo não clínico foi composto por indivíduos sem diagnósticos psiquiátricos especificados.

### Instrumentos

Os seguintes instrumentos foram utilizados neste estudo: ficha para levantamento sociodemográfico (e.g., sexo, idade, escolaridade) e o BDI-II (Beck et al., 1996), escala de autoavaliação que mensura a gravidade dos sintomas depressivos (nos últimos 15 dias) por meio de 21 itens com quatro respostas (variando de 0 a 3). A soma dos escores de todos os itens fornece uma pontuação máxima de 63. Escores mais elevados indicam gravidade

acentuada. No presente estudo, o BDI-II apresentou aceitável consistência interna (alfa de Cronbach = 0,91) e a correlação item-total variou de 0,46 a 0,66.

### Procedimentos de coleta de dados

Este estudo possui um delineamento transversal e quantitativo. A amostra foi obtida de maneira intencional (Coulacoglou & Saklofske, 2017) com aplicação presencial, efetuada por alunos de mestrado e iniciação científica, devidamente treinados. Os instrumentos foram aplicados em ambiente hospitalar (grupo psiquiátrico); em centros de atendimento especializados (grupo clínico); e em escolas públicas e privadas do ensino fundamental, universidades e grupos da terceira idade (grupo não clínico). A pesquisa foi aprovada pelo Comitê de Ética da Pontifícia Universidade Católica do Rio Grande do Sul (PUC/RS). Os procedimentos éticos para a pesquisa

com seres humanos foram seguidos (Conselho Nacional de Saúde, 1996).

### Procedimentos de análise de dados

O processo de validade de conteúdo do instrumento BDI-II para a população brasileira foi realizado em estudo anterior (Gorenstein et al., 2011). Assim sendo, na presente investigação, tais procedimentos não foram descritos.

No tocante à análise de dados, os dados descritivos da amostra ( $N=2.385$ ) foram inicialmente mensurados com o programa *Statistical Package for the Social Sciences* (SPSS, versão 22). Após, avaliou-se a normalidade multivariada por meio do programa FACTOR, versão 10 (Lorenzo-Seva & Ferrando, 2013). O índice de Mardia apresentou o valor de 26,7 ( $p \leq 0,05$ ), indicando uma amostra com distribuição não normal (índice de Mardia  $> 3,0, p \leq 0,05$ ; Mardia, 1970). Posteriormente, seguiram-se as seguintes etapas: (1) análise do modelo de mensuração – avalia a contribuição de cada item (indicador) em um construto (Hair et al., 2009); (2) teste de equivalência proposto por Yuan, Chan, Marcoulides e Bentler (2016) – identifica o limite tolerável da má especificação (*minimum tolerable size*, T-size) de um modelo; (3) avaliação da invariabilidade do modelo – verifica a equivalência fatorial entre diferentes grupos (Damásio, 2013; Hair et al., 2009); (4) estabelecimento de pontos de cortes para o BDI-II por meio da Análise de Classe Latente (e.g. Allan et al., 2014); e (5) verificação da variabilidade das classes latentes (Damásio, 2013; Hair et al., 2009) estabelecidas por meio da Análise de Classe Latente (Berlin et al., 2013; Muthén, 2008).

### Modelo de mensuração

Na ESEM, com o uso do programa Mplus (versão 7.11; Asparouhov & Muthén, 2009; Muthén, du Toit, & Spisic, 1997; Muthén & Muthén, 1998-2015), a estimação foi efetuada com o método Mínimos Quadrados Ponderados Robustos – média e variância ajustadas (*Robust Weighted Least Squares – Mean and Variance Adjusted*; WLSMV), apropriado para dados ordinais, tendo em vista que o BDI-II possui quatro pontos de respostas; e amostras não normais (índice de Mardia=26,7,  $p \leq 0,05$ ; Mardia, 1970). Examinou-se a possibilidade de quatro modelos: 1, 2, 3 fatores e o bifator. Para os modelos com u, dois ou três fatores, o corte das cargas fatoriais foi estipulado em 0,40 (Damásio, 2012). O modelo bifator foi analisado com base nos seguintes coeficientes (Dominguez-Lara & Rodriguez, 2017; Reise, 2012; Rodriguez et al., 2016): 1. ômega hierárquico ( $\omega_H$ ) – mensura a variância total atribuída ao fator geral; 2. de Hancock (HH.G) – informa quanto bem o conjunto de itens representa a variável latente; 3. Variância Comum Explicada (*Explained Common Variance*, ECV) – clarifica o montante de variância comum atribuída ao fator geral; 4. ECV-I – indica o percentual de variância explicado pelo fator geral de cada item relacionado aos subfatores; e 5.

Percentual de Correlações Não Contaminado (*Percentage of Uncontaminated Correlations*, PUC) – avalia o percentual de correlações não contaminado pela multidimensionalidade. Os valores desses índices devem ser  $\geq 0,70$ . Para os cálculos do  $\omega_H$ ,  $H_{H.G}$ , ECV, ECV-I e PUC, utilizou-se o módulo disponibilizado por Dominguez-Lara e Rodriguez (2017).

Além disso, os quatro modelos (1, 2, 3 fatores e bifator) foram avaliados por meio de indicadores de qualidade de modelo (medidas utilizadas para avaliar se o modelo possui uma especificação plausível). Os índices e os respectivos valores de referência, utilizados neste estudo, foram: Índice de Ajuste Comparativo (*Comparative Fit Index*, CFI  $> 0,95$ ), Índice de Tucker-Lewis (TLI  $> 0,95$ ), Raiz Quadrada da Média do Erro de Aproximação (*Root Mean Square Error of Approximation*, RMSEA  $< 0,06$ ; Intervalo de Confiança [IC], 90%; Hu & Bentler, 1998) e Raiz Quadrada Média Residual Padronizada (*Standardized Root Mean Square Residual*, SRMR  $< 0,08$ ; Hair et al., 2009).

### Teste de equivalência

Após a eleição do modelo mais plausível por meio da ESEM, pelo fato de amostras maiores tenderem a um valor significativo ( $p \leq 0,05$ ) do qui-quadrado ( $\chi^2$ ; Hair et al., 2009), efetuou-se o teste de equivalência proposto por Yuan et al. (2016). Nessa estratégia, os índices RMSEA e CFI são adotados para a avaliação da má especificação de um modelo (T-size). Os limites superiores dos intervalos de confiança (IC<sub>s</sub>) de ambos os índices (RMSEA e CFI) demonstram se a violação do modelo possa ser suportada; e a hipótese nula (H<sub>0</sub>), aceita. Essa estratégia permite que os modelos sejam classificados como: “excelentes”, “bons”, “aceitáveis”, “mediócras” e “pobres”. O *software R Studio* (versão *Windows*) foi utilizado para tais cálculos.

### Invariabilidade do modelo

Para verificar a invariabilidade do modelo, realizou-se a técnica AFCMG (Damásio, 2013; Hair et al., 2009). Para tanto, a amostra ( $N=2.385$ ) foi dividida pela técnica das metades partidas (grupo A,  $n=1.192$ ; grupo B,  $n=1.193$ ; Coulacoglou & Saklofske, 2017) com auxílio do programa SPSS. A comparação entre as amostras foi efetuada com o programa Mplus e o estimador WLSMV (Muthén & Muthén, 1998-2015).

Na primeira etapa, estimou-se os grupos separadamente com todos os parâmetros livres. Posteriormente, testou-se as seguintes invariabilidades: modelo 1, configural (todos os parâmetros livres); modelo 2, métrica (construção nas cargas fatoriais); modelo 3, escalar (construção nos interceptos); modelo 4, estrutural (construção nas médias); e modelo 5, residual (construção nos erros). Os grupos em separado e a invariância configural foram avaliados por meio dos seguintes índices: CFI ( $> 0,95$ ), RMSEA ( $< 0,06$ , CI90%) e TLI ( $< 0,95$ ; Hair et al., 2009; Hu & Bentler, 1998); bem como as demais

invariâncias com base nas diferenças do  $\chi^2$  ( $\Delta\chi^2, p > 0,05$ ), CFI ( $\Delta CFI < 0,01$ ) e TLI ( $\Delta TLI < 0,02$ ; Damásio, 2013).

### Pontos de corte

Com o auxílio do programa Mplus, a análise de classe latente foi efetuada com o método de máxima verossimilhança com erro-padrão e qui-quadrado robustos (MLR; Muthén, 2008). Comparou-se a possibilidade de duas, três e quatro classes. Para a verificação do melhor modelo de classes, utilizou-se os seguintes testes: (1) Lo-Mendell-Rubin (LMR; Lo, Mendell, & Rubin, 2001), (2) Critério de Informação Bayesiano (*Bayesian Information Criterion*, BIC; Schwarz, 1978) e (3) tamanho ajustado da amostra BIC (*sample-size adjusted BIC*, aBIC; Sclove, 1987). O teste LMR compara o modelo de classes proposto com outro com menos classes (e.g., 2 x 3), fornecendo determinado *p*-valor que indica se existe ou não diferença estatística entre as classes avaliadas, sendo preferíveis os modelos cujas as diferenças apresentem *p*-valores significativos ( $p \leq 0,05$ ; Berlin et al., 2013). Em relação aos testes BIC e aBIC, modelos com valores menores nesses testes podem ser considerados com melhor ajuste. Após a seleção do melhor modelo de classes, os pontos de corte foram estabelecidos por meio dos escores totais do BDI-II em relação aos indivíduos selecionados como pertencentes a cada classe (e.g. Allan et al., 2014).

### Variabilidade das classes latentes

Na sequência, verificou-se a possibilidade de variabilidade entre as classes latentes estabelecidas na análise de classe latente (Berlin et al., 2013; Muthén, 2008). Os procedimentos da AFCMG (Damásio, 2013; Hair et al., 2009) foram os mesmos utilizados na etapa da verificação da Invariabilidade do modelo. A possibilidade de diferença entre as classes foi avaliada por meio dos seguintes índices:  $\Delta\chi^2$  ( $p \leq 0,05$ ),  $\Delta CFI$  ( $> 0,01$ ) e  $\Delta TLI$  ( $> 0,02$ ; Damásio, 2013).

## Resultados

Em coerência com os objetivos do estudo e com os procedimentos de análise de dados descritos em seções anteriores do manuscrito, e na busca de aferir maior clareza na exposição dos achados do estudo, os resultados serão apresentados com determinada ordem. Inicialmente, serão demonstradas as análises quanto ao modelo de mensuração (Hair et al., 2009), seguidas do teste de equivalência/teste de hipótese (Yuan et al., 2016) e da avaliação da invariabilidade do modelo (Damásio, 2013; Hair et al., 2009). Na sequência, serão apresentados os pontos de cortes para o BDI-II, aferidos por meio da análise de classe latente (Allan et al., 2014), bem como serão explanados os achados no tocante à verificação da variabilidade das classes latentes (Berlin et al., 2013; Damásio, 2013; Hair et al., 2009; Muthén, 2008).

### Modelo de mensuração

Na ESEM, o modelo 1 (unidimensional) apresentou os índices CFI e TLI marginalmente abaixo dos limites adotados para este estudo (CFI e TLI  $> 0,95$ ; Hair et al., 2009; Hu & Bentler, 1998; Tabela 2); bem como alguns itens (2, 5, 8, 10, 11, 13, 16, 17, 18, 19, 20, 21) exibiram cargas fatoriais abaixo do recomendado ( $\geq 0,70$ ; Hair et al. 2009), aspecto não apropriado para modelos unidimensionais (Reise, 2012). Já o modelo 2 (dois fatores) evidenciou certos itens (12, 17, 18 e 21) que não se discriminaram entre os fatores, ou seja, não revelaram cargas fatoriais específicas a um único fator. O modelo com três fatores apresentou todas as cargas fatoriais abaixo de 0,40 (Damásio, 2012), assim, não foi descrito neste estudo. Portanto, os modelos com um, dois e três fatores não puderam ser considerados como plausíveis.

Tabela 2  
Modelos Fatoriais do BDI-II

Itens	Modelo 1		Modelo 2		Modelo 3	
	Fator 1	Fator 1	Fator 2	Fator G	Fator 1	Fator 2
1. Tristeza	<b>0,76*</b>	<b>0,73*</b>		<b>0,72*</b>	0,27*	
2. Pessimismo	<b>0,68*</b>	<b>0,56*</b>		<b>0,64*</b>	0,23*	
3. Fracasso passado	<b>0,72*</b>	<b>0,70*</b>		<b>0,73*</b>	0,34*	
4. Perda de prazer	<b>0,78*</b>	<b>0,53*</b>		<b>0,81*</b>		
5. Sentimentos de culpa	<b>0,69*</b>	<b>0,78*</b>		<b>0,70*</b>	0,34*	
6. Sentimentos de punição	<b>0,70*</b>	<b>0,84*</b>		<b>0,71*</b>	0,34*	
7. Autoestima	<b>0,79*</b>	<b>0,83*</b>		<b>0,71*</b>	0,41*	
8. Autocrítica	<b>0,64*</b>	<b>0,70*</b>		<b>0,57*</b>	0,36*	
9. Pensamentos ou desejos suicidas	<b>0,71*</b>	<b>0,64*</b>		<b>0,68*</b>	0,25*	
10. Choro	<b>0,67*</b>	<b>0,58*</b>		<b>0,65*</b>	0,19*	
11. Agitação	<b>0,63*</b>	<b>0,51*</b>		<b>0,65*</b>		
12. Perda de interesse	<b>0,70*</b>	0,40*	0,38*	<b>0,73*</b>		
13. Indecisão	<b>0,65*</b>	<b>0,54*</b>		<b>0,62*</b>	0,21*	

Tabela 2 (continuação)  
Modelos Fatoriais do BDI-II

Itens	Modelo 1	Modelo 2		Modelo 3		
	Fator 1	Fator 1	Fator 2	Fator G	Fator 1	Fator 2
14. Desvalorização	<b>0,77*</b>	<b>0,66*</b>		<b>0,71*</b>	0,30*	
15. Falta de energia	<b>0,70*</b>		<b>0,75*</b>	<b>0,68*</b>		0,41*
16. Alterações no padrão de sono	<b>0,60*</b>		<b>0,43*</b>	<b>0,60*</b>		0,15*
17. Irritabilidade	<b>0,65*</b>	0,36*	0,35*	<b>0,67*</b>		
18. Alterações no apetite	<b>0,62*</b>	0,36*	0,32*	<b>0,63*</b>		0,10*
19. Dificuldade de concentração	<b>0,66*</b>		<b>0,48*</b>	<b>0,67*</b>		0,17*
20. Cansaço ou fadiga	<b>0,68*</b>		<b>0,80*</b>	<b>0,75*</b>		0,42*
21. Perda de interesse por sexo	<b>0,64*</b>	0,36*	0,35*	<b>0,67*</b>		
Covariância F1-F2			<b>0,68*</b>			
Índices	Modelo 1	Modelo 2		Modelo 3		
$\chi^2$	1.838,14	1.024,13		968,90		
GL	189	169		173		
p-valor	<0,01	<0,01		<0,01		
CFI	0,95	0,97		0,98		
TLI	0,94	0,97		0,97		
RMSEA	0,05	0,05		0,04		
(IC90%)	(0,05-0,06)	(0,04-0,05)		(0,04-0,05)		
SRMR	0,05	0,04		0,04		

Nota. Negritos=cargas fatoriais acima de 0,40; Valores omitidos nos modelos 1 e 2=cargas fatoriais abaixo de 0,30; Valores omitidos no modelo 3=cargas fatoriais não significativas ( $p>0,05$ ); CFI=Índice de Ajuste Comparativo; Fator G=fator geral; GL=graus de liberdade; RMSEA=Raiz Quadrada da Média do Erro de Aproximação; SRMR=Raiz Quadrada Média Residual Padronizada; TLI=Índice de Tucker-Lewis;  $\chi^2$ =qui-quadrado. \*  $p\leq 0,05$

Na sequência, testou-se o modelo bifator (modelo 3), encontrando um fator geral oblíquo (Depressão) e dois subfatores oblíquos (Cognitivo [fator 1] e Somático [fator 2]). Os itens 4, 11, 12, 17 e 21 não apresentaram cargas fatoriais significativas ( $p>0,05$ ) nos subfatores oblíquos (Cognitivo e Somático), permanecendo relacionados apenas ao fator geral (Depressão). Esse modelo alcançou índices de precisão aceitáveis (ECV=0,84, PUC=0,76,  $\omega_H=0,86$  e  $H_{H,G}=0,93$ ; Dominguez-Lara & Rodriguez, 2017; Reise, 2012; Rodriguez et al., 2016); bem como os valores do ECV-I foram maiores que 0,70 para todos os itens relacionados aos subfatores. Ademais, os índices de qualidade do modelo indicaram a possibilidade do modelo 3 ser plausível (CFI>0,95, TLI>0,95, RMSEA<0,06 e SRMR<0,08; Hair et al., 2009; Hu & Bentler, 1998).

### Teste de hipótese

Em relação ao modelo bifator confirmatório, o teste  $\chi^2$  (968,90) apresentou-se significativo ( $p\leq 0,05$ ); e a razão  $\chi^2/GL$  (5,60), inaceitável ( $>3$ ; Hair et al., 2009). Sendo assim, efetuou-se o teste de hipótese ( $H_0$ ) com base na estratégia de Yuan et al. (2016). Para o modelo 3 (bifator), os valores de T-size dos índices RMSEA e CFI foram, respectivamente, 0,054 (IC: 0,021-0,057) e 0,970 (IC: 0,943-0,987; modelo independente = 32.955,10). Tais valores classificaram o modelo como "bom".

### Invariabilidade do modelo

Na verificação da invariabilidade do modelo (Tabela 3), os grupos A ( $n=1.192$ ) e B ( $n=1.193$ ) revelaram índices de qualidade do modelo de acordo com os valores recomendados (CFI>0,95; TLI>0,95; RMSEA<0,06; Hair et al., 2009; Hu & Bentler, 1998). O mesmo fenômeno pôde ser observado em relação ao modelo 1, configural (parâmetros livres).

Além disso, as diferenças entre modelos (1 a 5;  $\Delta\chi^2$ ,  $p>0,05$ ;  $\Delta CFI<0,01$ ;  $\Delta TLI<0,02$ ; Damásio, 2013) evidenciaram a viabilidade de invariabilidade métrica (modelo 2), escalar (modelo 3), estrutural (modelo 4) e residual (modelo 5). Esses dados sugeriram a possibilidade do BDI-II apresentar equivalência fatorial para a amostra avaliada (Coulacoglou & Saklofske, 2017; Damásio, 2013; Hair et al., 2009).

### Pontos de corte

Por meio da análise de classe latente, observou-se que o modelo com quatro classes (e.g. Beck et al., 1996) não pôde ser suportado, pois não houve diferença estatisticamente significativa entre 3 e 4 classes (LMR=966,18,  $p>0,05$ ). Houve diferenças estatisticamente significativas entre os modelos com uma e duas classes (LMR=10.297,85;  $p\leq 0,05$ ), bem como entre os modelos com duas e três classes (LMR=2.623,76;  $p\leq 0,05$ ). Contudo, os valores de BIC e aBIC foram

menores para o modelo com 3 classes (BIC=81.182,70, aBIC=80.575,86) em comparação ao modelo com duas classes (BIC=83.313,42, aBIC=82.909,91), sugerindo que o modelo com três classes possa ser considerado como preferível (Berlin et al., 2013).

No modelo com três classes, a classe 1 ( $n=1.377$ , 57,74%) caracterizou-se pela maioria de respostas 0 e 1 ( $M_r=4,05$ ,  $DP_r=3,22$ ); a classe 2 ( $n=617$ , 25,87%), respostas 1 e 2 ( $M_r=14,38$ ,  $DP_r=2,75$ ); a classe 3 ( $n=391$ , 16,39%), respostas 2 e 3 ( $M_r=30,78$ ,  $DP_r=7,78$ ). Assim, as classes foram categorizadas como: sintomas depressivos mínimos (classe 1), leves (classe 2) e moderados a

graves (classe 3). Os pontos de corte correspondentes foram: 0 a 12, sintomas mínimos; 13 a 21, leves; e acima de 22, moderados a graves. Os pontos de corte encontrados neste estudo não corresponderam aos sugeridos por Beck et al. (1996; sintomas depressivos mínimos, 0-13; leves, 14-19; moderados, 20-28; e graves, 29-63). Com base nos pontos de corte propostos por Beck et al. (1996) e em relação à amostra da região metropolitana de Porto Alegre, Rio Grande do Sul, Gorenstein et al. (2011) observaram que 68,4% da amostra apresentaram sintomas depressivos mínimos; 12,3%, leves; 11,0%, moderados; e 8,3% revelaram sintomas depressivos graves.

Tabela 3  
Invariabilidade do Modelo

Invariabilidade		Índices						Diferenças				
		$\chi^2$	GL	$p$	CFI	TLI	RMSEA	$\Delta\chi^2$	$\Delta$ GL	$p$	$\Delta$ CFI	$\Delta$ TLI
Grupos	A	608,19	174	<0,01	0,97	0,97	0,05					
	B	600,17	174	<0,01	0,97	0,96	0,05					
Modelos	1	1.208,36	348	<0,01	0,97	0,96	0,05					
	2	1.249,85	381					41,49	33	0,15	0,001	0,006
	3	1.280,68	402					30,83	21	0,08	0,001	0,006
	4	1.354,26	459					73,58	57	0,07	0,001	0,006
	5	1.452,84	537					98,58	78	0,06	0,004	0,008

Nota. CFI=Índice de Ajuste Comparativo; GL=graus de liberdade;  $p$ =nível de significância; RMSEA=Raiz Quadrada da Média do Erro de Aproximação; TLI=Índice de Tucker-Lewis;  $\chi^2$ =qui-quadrado;  $\Delta$ CFI=diferença do índice CFI;  $\Delta$ GL= diferença dos graus de liberdade;  $\Delta$ TLI=diferença do índice TLI;  $\Delta\chi^2$ =diferença do  $\chi^2$ .

Na avaliação da variabilidade das classes por meio da técnica AFCMG (Tabela 4), a classe 1 (sintomas depressivos mínimos), 2 (sintomas depressivos leves) e 3 (sintomas depressivos moderados/graves), bem como a invariabilidade configural alcançaram valores dos índices de qualidade do modelo aceitáveis (CFI>0,95; TLI>0,95; RMSEA<0,06; Damásio, 2013; Hair et al., 2009; Hu &

Bentler, 1998). Além disso, as diferenças ( $\Delta\chi^2$ ,  $p\leq 0,05$ ;  $\Delta$ CFI>0,01;  $\Delta$ TLI>0,02; Damásio, 2013) entre os modelos avaliados (2, métrica; 3, escalar; 4, estrutural; e 5, residual; Damásio, 2013; Hair et al., 2009) indicaram a possibilidade das três classes serem diferenciadas em termos da gravidade dos sintomas depressivos (mínimos, leves e moderados/graves).

Tabela 4  
Diferença entre Três Classes de Gravidade dos Sintomas Depressivos

Variâncias		Índices						Diferenças				
		$\chi^2$	GL	$p$	CFI	TLI	RMSEA	$\Delta\chi^2$	$\Delta$ GL	$p$	$\Delta$ CFI	$\Delta$ TLI
Classes	1	664,72	200	<0,01	0,97	0,96	0,04					
	2	432,60	200	<0,01	0,97	0,96	0,04					
	3	351,04	200	<0,01	0,97	0,96	0,04					
Modelos	1	1.448,36	599	<0,01	0,98	0,97	0,02					
	2	1.612,15	633	<0,01				163,80	34	<0,01	0,06	0,06
	3	1.956,12	654	<0,01				343,97	21	<0,01	0,08	0,07
	4	2.857,19	657	<0,01				901,07	3	<0,01	0,13	0,12
	5	4.419,99	678	<0,01				1.562,80	21	<0,01	0,61	0,62

Nota. CFI=Índice de Ajuste Comparativo; GL=graus de liberdade;  $p$ =nível de significância; RMSEA=Raiz Quadrada da Média do Erro de Aproximação; TLI=Índice de Tucker-Lewis;  $\chi^2$ =qui-quadrado;  $\Delta$ CFI=diferença do índice CFI;  $\Delta$ GL= diferença dos graus de liberdade;  $\Delta$ TLI=diferença do índice TLI;  $\Delta\chi^2$ =diferença do  $\chi^2$ .

## Discussão

A presente investigação buscou avaliar as propriedades psicométricas do BDI-II em uma amostra de indivíduos da região metropolitana de Porto Alegre, Rio Grande do Sul, Brasil. Assim, por meio de técnicas da Teoria do Traço Latente, soluções fatoriais e pontos de corte para o instrumento foram explorados, bem como foi verificada a possibilidade de invariabilidade e variância entre grupos.

No que diz respeito às soluções fatoriais, os dados suportaram o modelo bifator confirmatório com um fator geral oblíquo (Depressão) e dois subfatores oblíquos (Cognitivo e Somático) não correlacionados. Investigações prévias apoiaram esses achados (Brouwer et al., 2013; Dere et al., 2015; Ward, 2006). Ademais, o modelo bifator, analisado neste estudo, revelou satisfatórias medidas de precisão, significando que a maior parte da variância entre os itens ocorreu especialmente pela presença de um fator geral, ou seja, houve indicativos de que o instrumento possa ser predominantemente unidimensional (Dominguez-Lara & Rodriguez, 2017; Reise, 2012; Rodriguez et al., 2016). De acordo com Bringmann et al. (2015), os sintomas depressivos descritos no BDI-II formam uma rede de interações cuja medida principal pode ser representada por um construto único (Depressão).

Não obstante, o modelo bifator confirmatório, avaliado neste estudo, alcançou um  $p$ -valor significativo ( $\leq 0,05$ ) e uma relação  $\chi^2/GL$  inaceitável ( $> 3$ ; Hair et al., 2009). Esse fenômeno pode ser explicado em decorrência do tamanho amostral ( $N=2.385$ ) em relação ao número de parâmetros a serem estimados (58). Tais achados também foram observados por Strunk e Lane (2017) em relação a estudos sobre o BDI-II.

Em atenção ao  $p$ -valor significativo, foi realizado o teste de hipótese proposto por Yuan et al. (2016), que revelou que o modelo bifator pôde ser classificado como "bom". Portanto, a violação do modelo pôde ser suportada; e a hipótese nula ( $H_0$ ), aceita; ou seja, o modelo pôde ser considerado como próximo de uma base populacional.

Também, neste estudo, observou-se a perspectiva do BDI-II apresentar invariância amostral referente ao modelo bifator confirmatório proposto. Esses resultados, em conjunto com os achados de outras investigações (Dere et al., 2015; Strunk & Lane, 2017; Whisman et al., 2013), sugeriram a existência de evidências de que o instrumento possa apresentar invariabilidade de medida, ou seja, precisão para a mensuração da gravidade dos sintomas depressivos (Coulacoglou & Saklofske, 2017).

Ademais, neste estudo, foi encontrada a possibilidade de três classes latentes distintas em termos da gravidade dos sintomas depressivos (mínimos, leves e moderados/graves) para a amostra avaliada. Esses resultados não foram apoiados por outros estudos que identificaram

quatro classes de sintomas depressivos (mínimo, leve, moderado e grave; Beck et al., 1996; Gorenstein et al., 2011) ou mesmo duas classes (com e sem sintomas depressivos; Kjaergaard et al., 2014; Plourde et al., 2016).

Também, este estudo verificou que em torno de 16% da amostra apresentou sintomas depressivos situados nos níveis moderado a grave, utilizando o ponto de corte de 22 pontos. Por sua vez, com base em amostra da região metropolitana de Porto Alegre, Rio Grande do Sul, o estudo de Gorenstein et al. (2011) encontrou 18% da amostra com sintomas depressivos de moderado (11%) a grave (8%) com o ponto de corte de 20 pontos. Diferentes pontos de corte tendem a alterar a prevalência dos sintomas depressivos, mesmo se tratando de amostras da mesma localidade (Hutz, Bandeira, & Trentini, 2015).

Além disso, o uso de técnicas distintas para a aferição dos pontos de corte em relação ao BDI-II pode ser uma das explicações para achados dessemelhantes entre estudos. Por exemplo, na versão do BDI-II adaptada para a população brasileira, foi utilizada a análise discriminante, comparando os pontos de corte propostos por Beck et al. (1996) com a Entrevista Clínica Estruturada baseada no DSM-IV (Gorenstein et al., 2011). Este estudo valeu-se da análise de classe latente para obtenção de cortes ao BDI-II. De acordo com Allan et al. (2014), a técnica pode ser útil para a proposição de cortes para instrumentos psicológicos na medida em que estabelece grupos a partir das respostas de indivíduos que compartilham padrões de dados semelhantes.

Além do mais, destaca-se que a prevalência dos sintomas depressivos pode se diferenciar de acordo com a cultura, sexo, escolaridade, renda, idade, entre outros fatores (Argimon et al., 2016). Por exemplo, a amostra (região metropolitana de Porto Alegre Porto Alegre/RS) desta investigação contou com 54% de mulheres. No estudo de Gorenstein et al. (2011), 62% da amostra (região metropolitana da cidade de São Paulo/SP e de Porto Alegre/RS) eram de mulheres. Estudos prévios demonstraram que a prevalência de sintomas depressivos pode ser maior em mulheres (Silva et al., 2014; *World Health Organization*, 2017; Wu & Huang, 2014); o que também pode justificar prevalências diferenciadas dos sintomas depressivos entre estudos, dentre outros aspectos.

Em termos de significância clínica, o achado deste estudo, da possibilidade de três classes de sintomas depressivos, pode indicar que não exista uma distinção clara entre a gravidade e a percepção subjetiva dos sintomas depressivos moderados e graves para a população avaliada. Segundo Muthén e Muthén (1998-2015), a análise de classe latente pode oferecer vantagens sobre as técnicas psicométricas tradicionais já que é um método centrado na pessoa e não na relação entre variáveis. Portanto, um dos benefícios da proposição de três classes de sintomas depressivos, encontradas neste estudo, recai na questão de que o método utilizado (análise de classe

latente) respalda-se na percepção dos sujeitos, tendo em vista que o BDI-II é uma escala de autorrelato; enquanto outros estudos (e.g., Beck et al., 1996; Kjaergaard et al., 2014; Plourde et al., 2016) basearam-se na avaliação de especialistas (Entrevista Estruturada com base em versões do DSM) para diagnóstico da gravidade dos sintomas depressivos comparando-a com os escores do BDI-II. Contudo, para a área clínica, a ausência da detecção mais acurada dos casos com sintomas depressivos graves pode comprometer as terapêuticas necessárias. Porém, o BDI-II pode ser considerado um instrumento de triagem, assim, não sendo capaz de distinguir os diferentes tipos de síndromes depressivas (Ely et al., 2014; Wang & Gorenstein, 2013a). Segundo Wang & Gorenstein, o profissional deve buscar evidências complementares para a interpretação dos escores do BDI-II para tomar decisões clínicas. Para a área da pesquisa, a utilização de três classes de sintomas depressivos, em detrimento de quatro classes, pode ou não comprometer os resultados dependendo dos objetivos propostos e métodos utilizados (e.g., estudos correlacionais). Destaca-se, assim, que investigações visando avaliar a validade concorrente (comparação entre instrumentos) são imprescindíveis para possibilitar que a sensibilidade e especificidade dos pontos de corte encontrados neste estudo sejam verificadas, reduzindo os casos de falso positivos e negativos (Hutz et al., 2015).

Em todo o caso, como alegou Strunk e Lane (2017), as técnicas baseadas na Teoria do Traço Latente não são suficientes para determinar a acurácia diagnóstica (validade de medida) do BDI-II, limitando-se a apenas sugerir melhores soluções. Tal questão representa uma das principais limitações deste estudo. Embora o uso da análise de classe latente seja um aspecto relevante desta investigação, pois a técnica permite que classes sejam identificadas com base nas diferenças em termos da variabilidade amostral, não foram localizadas outras investigações que tenham utilizado essa técnica com o BDI-II. Também, como limitação, a presente investigação valeu-se apenas de algumas técnicas pertencentes à TTL (e.g., ESEM) de modo a explorar as capacidades psicométricas do BDI-II; todavia, futuros estudos poderão ampliar os conhecimentos sobre o instrumento abarcando outros procedimentos da psicometria moderna, tal como a Teoria da Resposta ao Item (TRI; Pasquali, 2009). Igualmente, não

foram aplicados outros instrumentos que permitissem que a análise concorrente pudesse ser efetuada. Assim sendo, os resultados deste estudo não possuem base de comparação, o que imediatamente sugere que os pontos de corte identificados não devem ser utilizados na prática clínica antes da realização de mais investigações a respeito. Ademais, vale lembrar que a amostra deste estudo foi oriunda da região metropolitana de Porto Alegre, Rio Grande do Sul, Brasil. Portanto, pode ser considerada como uma amostra com determinada homogeneidade cultural, sendo outra limitação deste estudo, já que comparações com outras amostras não foram efetuadas.

Portanto, a partir deste estudo, evidências foram encontradas de que o BDI-II possa ser prioritariamente unidimensional e apresentar precisão de medida para a mensuração da gravidade dos sintomas depressivos. Também, pode ser possível que o instrumento suporte três distintas classes em termos de gravidade dos sintomas depressivos (mínimos, leves e moderados a graves). Certamente, mais investigações sobre as propriedades psicométricas do BDI-II são recomendadas.

### Considerações finais

Este estudo avaliou as propriedades psicométricas do BDI-II com base em algumas técnicas da Teoria do Traço Latente. Assim, os dados apresentados nesse estudo forneceram subsídios para concluir que um modelo unidimensional com subfatores mais restritos possa melhor se adequar ao instrumento. Ademais, os dados apoiaram a invariabilidade da medida, o que sugeriu que o BDI-II possa ser aplicado em diferentes amostras. Também os dados indicaram que o instrumento possa contar com três classes de gravidade de sintomas depressivos (mínimos, leves e moderados-severos). Porém, em termos de acurácia diagnóstica do instrumento, as técnicas utilizadas não podem ser consideradas suficientes para conclusões definitivas. Nesse sentido, torna-se importante a realização de outros estudos, contando com diferentes amostras (não clínicas e, especialmente, clínicas) para avaliar se o modelo preponderantemente unidimensional (bifator) possa ser suportado; e, especialmente, buscar indicativos de sensibilidade e especificidade dos pontos de corte verificados a partir desta investigação.

### Referências

- Allan, N. P., Raines, A. M., Capron, D. W., Norr, A. M., Zvolensky, M. J., & Schmidt, N. B. (2014). Identification of anxiety sensitivity classes and clinical cut-scores in a sample of adult smokers: Results from a factor mixture model. *Journal of Anxiety Disorders*, 28(7), 696-703. doi: 10.1016/j.janxdis.2014.07.006
- Argimon, I. I. L., Paloski, L. H., Farina, M., & Irigaray, T. Q. (2016). Aplicabilidade do Inventário de Depressão de Beck-II em idosos: Uma revisão sistemática. *Avaliação Psicológica*, 15(spe), 11-17. Recuperado de [http://pepsic.bvsalud.org/scielo.php?script=sci\\_arttext&pid=S1677-04712016000300003&lng=pt&tlng=pt](http://pepsic.bvsalud.org/scielo.php?script=sci_arttext&pid=S1677-04712016000300003&lng=pt&tlng=pt)
- Asparouhov, T., & Muthén, B. (2009). Exploratory structural equation modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, 16(3), 397-438. doi: 10.1080/10705510903008204

- Associação Americana de Psiquiatria. (1995). *Manual diagnóstico e estatístico de transtornos mentais: DSM-IV®*. 4a ed. Porto Alegre, RS: Artes Médicas.
- Associação Americana de Psiquiatria. (2014). *Manual diagnóstico e estatístico de transtornos mentais: DSM-V®*. Porto Alegre, RS: Artes Médicas.
- Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Manual for the Beck Depression Inventory-II*. San Antonio, TX: Psychological Corporation.
- Beck, A. T., Ward, C. H., Mendelson, M., Moch, J., & Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry*, 4(6), 561-571. doi:10.1001/archpsyc.1961.01710120031004
- Berlin, K. S., Williams, N. A., & Parra, G. R. (2013). An introduction to latent variable mixture modeling (Part 1): Overview and cross-sectional latent class and latent profile analyses. *Journal of Pediatric Psychology*, 39(2), 174-187. doi: 10.1093/jpepsy/jst084
- Bhardwaj, M., Price, J., Landry, M., Harvey, P., & Hensel, J. M. (no prelo). Association Between severity of depression and cardiac risk factors among women referred to a cardiac rehabilitation and prevention clinic. *Journal of Cardiopulmonary Rehabilitation and Prevention*, 38(5), 291-296. doi: 10.1097/HCR.0000000000000311
- Bringmann, L. F., Lemmens, L. H., Huibers, M. J., Borsboom, D., & Tuerlinckx, F. (2015). Revealing the dynamic network structure of the Beck Depression Inventory-II. *Psychological Medicine*, 45(4), 747-757. doi: 10.1017/S0033291714001809
- Brouwer, D., Meijer, R. R., & Zevalkink, J. (2013). On the factor structure of the Beck Depression Inventory-II: G is the key. *Psychological Assessment*, 25(1), 136-145. doi: 10.1037/a0029228
- Coulacoglou, C., & Saklofske, D. H. (2017). *Psychometrics and psychological assessment: Principles and applications*. Cambridge, MA: Academic Press.
- Cowen, P. (2017). Depression: Symptomatology, diagnosis, and classification. Em C. Foster, & J. Herring (Eds.), *Depression: Law and ethics* (pp. 12-27). Oxford, UK: Oxford University Press.
- Cunha, J. A. (2001). *Manual da versão em português das escalas Beck*. São Paulo, SP: Casa do Psicólogo.
- Damásio, B. F. (2012). Uso da análise fatorial exploratória em psicologia. *Avaliação Psicológica*, 11(2), 213-228. Recuperado de [http://pepsic.bvsalud.org/scielo.php?script=sci\\_arttext&pid=S1677-04712012000200007&lng=pt&tlng=pt](http://pepsic.bvsalud.org/scielo.php?script=sci_arttext&pid=S1677-04712012000200007&lng=pt&tlng=pt)
- Damásio, B. F. (2013). Contribuições da análise fatorial confirmatória multigrupo (AFCMG) na avaliação de invariância de instrumentos psicométricos. *Psico-USF*, 18(2), 211-220. doi: 10.1590/S1413-82712013000200005
- Dere, J., Watters, C. A., Yu, S. C., Bagby, R. M., Ryder, A. G., & Harkness, K. L. (2015). Cross-cultural examination of measurement invariance of the Beck Depression Inventory-II. *Psychological Assessment*, 27(1), 68-81. doi: 10.1037/pas0000026
- Dominguez-Lara, S. A., & Rodriguez, A. (2017). Índices estadísticos de modelos bifactor. *Interacciones: Revista de Avances en Psicología*, 3(2), 59-65. doi: 10.24016/2017.v3n2.51
- Doshi, S., Keilp, J. G., Strobino, B., Mcelhiney, M., Rabkin, J., & Fallon, B. A. (2018). Depressive symptoms and suicidal ideation among symptomatic patients with a history of lyme disease vs two comparison groups. *Psychosomatics*, 59(5), 481-489. doi: 10.1016/j.psym.2018.02.004
- Ely, P., Nunes, M. F. O., & Carvalho, L. F. (2014). Avaliação psicológica da depressão: Levantamento de testes expressivos e autorrelato no Brasil. *Avaliação Psicológica*, 13(3), 419-426. Recuperado de [http://pepsic.bvsalud.org/scielo.php?script=sci\\_arttext&pid=S1677-04712014000300014&lng=pt&tlng=pt](http://pepsic.bvsalud.org/scielo.php?script=sci_arttext&pid=S1677-04712014000300014&lng=pt&tlng=pt)
- Gorenstein, C., Pang, W. Y., Argimon, I. L., & Werlang, B. S. G. (2011). Inventário Beck de Depressão-II. Manual. São Paulo, SP: Casa do Psicólogo.
- Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., & Tatham, R. L. (2009). *Análise Multivariada de Dados*. 6a ed. Porto Alegre, RS: Bookman.
- Harris, C. A., & D'Eon, J. L. (2008). Psychometric properties of the Beck Depression Inventory - Second Edition (BDI-II) in individuals with chronic pain. *Pain*, 137(3), 609-622. doi: 10.1016/j.pain.2007.10.022
- Hayes, S., S., Stoeckel, N., Napolitano, M. A., Collins, C., Wood, G. C., Seiler, J.,... Still, C. D. (2015). Examination of the Beck Depression Inventory-II Factor Structure among bariatric surgery candidates. *Obesity Surgery*, 25(7), 1155-1160. doi: 10.1007/s11695-014-1506-5
- Hu, L., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods*, 3(4), 424-453. doi: 10.1037/1082-989x.3.4.424
- Hutz, C. S., Bandeira, D. R., & Trentini, C. (2015). *Psicométrica*. Porto Alegre: Artmed, 2015.
- Kjaergaard, M., Arfwedson Wang, C. E., Waterloo, K., & Jorde, R. (2014). A study of the psychometric properties of the Beck Depression Inventory-II, the Montgomery and Åsberg Depression Rating Scale, and the Hospital Anxiety and Depression Scale in a sample from a healthy population. *Scandinavian Journal of Psychology*, 55(1), 83-89. doi: 10.1111/sjop.12090
- Lam, R. W. (2018). *Depression*. 3th ed. Oxford, UK: Oxford University Press.
- Lo, Y., Mendell, N. R., & Rubin, D. B. (2001). Testing the number of components in a normal mixture. *Biometrika*, 88(3), 767-778. doi: 10.1093/biomet/88.3.767
- Lorenzo-Seva, U., & Ferrando, P. J. (2013). FACTOR: A computer program to fit the Exploratory Factor Analysis model. *Behavior Research Methods*, 38(1), 88-91. doi: 10.3758/BF03192753
- Mardia, K. V. (1970). Measures of multivariate skewness and kurtosis with applications. *Biometrika*, 57(3), 519-530. doi: 10.1093/biomet/57.3.519
- Minas, H. (2017). Depression in the developing world. Em C. Foster, & J. Herring (Eds.), *Depression: Law and Ethics* (pp. 186-198). Oxford, UK: Oxford University Press.
- Muthén, B. O. (2008). Latent variable hybrids: Overview of old and new hybrids. Em G. R. Hancock & K. M. Samuelsen (Eds.), *Advances in Latent Variable Mixture Models* (pp. 1-24). Charlotte, NC: Information Age Publishing.
- Muthén, B. O., du Toit, S. H. C., & Spisic, D. (1997). *Robust inference using weighted least squares and quadratic estimating equations in latent variable modeling with categorical and continuous outcomes*. (Technical report). Los Angeles, CA: University of California.
- Muthén, L. K., & Muthén, B. O. (1998-2015). *Mplus User's Guide*. 8th ed. Los Angeles, CA: Muthén & Muthén. Recuperado de [https://www.statmodel.com/download/usersguide/MplusUserGuideVer\\_8.pdf](https://www.statmodel.com/download/usersguide/MplusUserGuideVer_8.pdf)
- Pasquali, L. (2009). Psicométrica. *Revista da Escola de Enfermagem da USP*, 43(spe), 992-999. doi: 10.1590/S0080-62342009000500002
- Plourde, A., Moullec, G., Bacon, S. L., Suartha, E., & Lavoie, K. L. (2016). Optimizing screening for depression among adults with asthma. *Journal of Asthma*, 53(7), 736-743. doi: 10.3109/02770903.2016.1145692
- Poole, H., Bramwell, R., & Murphy, P. (2009). The utility of the Beck Depression Inventory Fast Screen (BDI-FS) in a pain clinic population. *European Journal of Pain*, 13(8), 865-869. doi: 10.1016/j.ejpain.2008.09.017

- Primi, R. (2010). Avaliação psicológica no Brasil: Fundamentos, situação atual e direções para o futuro. *Psicologia: Teoria e Pesquisa*, 26(spe), 25-35. doi: 10.1590/S0102-37722010000500003
- Reise, S. P. (2012). The rediscovery of bifactor measurement models. *Multivariate Behavioral Research*, 47(5), 667-696. doi: 10.1080/00273171.2012.715555
- Resolução nº 196 (1996). Brasília, DF: Conselho Nacional de Saúde. Recuperado de [conselho.saude.gov.br/docs/Reso196.doc](http://conselho.saude.gov.br/docs/Reso196.doc).
- Rodriguez, A., Reise, S. P., & Haviland, M. G. (2016). Evaluating bifactor models: Calculating and interpreting statistical indices. *Psychological Methods*, 21(2), 137-150. doi: 10.1037/met0000045
- Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, 6(2), 461-464. Recuperado de <http://qwone.com/~jason/trg/papers/schwarz-dimension-78.pdf>
- Sclove, L. S. (1987). Application of model-selection criteria to some problems in multivariate analysis. *Psychometrika*, 52(3), 333-343. doi: 10.1007/BF02294360
- Silva, M. T., Galvao, T. F., Martins, S. S., & Pereira, M. G. (2014). Prevalence of depression morbidity among Brazilian adults: A systematic review and meta-analysis. *Revista Brasileira de Psiquiatria*, 36(3), 262-270. doi: 10.1590/1516-4446-2013-1294
- Strunk, K. K., & Lane, F. C. (2017). The Beck Depression Inventory, second edition (BDI-II): A cross-sample structural analysis. *Measurement and Evaluation in Counseling and Development*, 50(1-2: 50th Anniversary), 3-17. doi: 10.1177/0748175616664010
- Wang, Y.-P., & Gorenstein, C. (2013a). Assessment of depression in medical patients: A systematic review of the utility of the Beck Depression Inventory-II. *Clinics*, 68(9), 1274-1287. doi: 10.6061/clinics/2013(09)15
- Wang, Y.-P., & Gorenstein, C. (2013b). Psychometric properties of the Beck Depression Inventory-II: A comprehensive review. *Revista Brasileira de Psiquiatria*, 35(4), 416-431. doi: 10.1590/1516-4446-2012-1048
- Ward, L. C. (2006). Comparison of factor structure models for the Beck Depression Inventory-II. *Psychological Assessment*, 18(1), 81-88. doi: 10.1037/1040-3590.18.1.81
- Whisman, M. A., Judd, C. M., Whiteford, N. T., & Gelhorn, H. L. (2013). Measurement invariance of the Beck Depression Inventory – Second Edition (BDI-II) across gender, race, and ethnicity in College Students. *Assessment*, 20(4), 419-428. doi: 10.1177/1073191112460273
- World Health Organization. (2017). Depression and other common mental disorders. Global health estimates. Geneva, SU: WHO. Recuperado de <http://apps.who.int/iris/bitstream/handle/10665/254610/WHO-MSD-MER-2017.2-eng.pdf;jsessionid=3B6D9D7D9F25053FBD2426E3BED85EF9?sequencehttp://apps.who.int/iris/bitstream/10665/254610/1/WHO-MSD-MER-2017.2-eng.pdf>
- Wu, P. C., & Huang, T. W. (2014). Gender-related invariance of the Beck Depression Inventory II for Taiwanese adolescent samples. *Assessment*, 21(2), 218-226. doi: 10.1177/1073191112441243
- Yuan, K.-H., Chan, W., Marcoulides, G. A., & Bentler, P. M. (2016). Assessing structural equation models by equivalence testing with adjusted fit indexes. *Structural Equation Modeling: A Multidisciplinary Journal*, 23(3), 319-330. doi: 10.1080/10705511.2015.1065414

recebido em novembro de 2017  
aceito em maio de 2018

## Sobre os autores

**Marli Appel da Silva** é psicóloga, doutora em Psicologia pela Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS), especialista em Avaliação Psicológica (PUCRS) e colaboradora Técnica do Grupo de Pesquisa Avaliação e Intervenção no Ciclo Vital (AICV).

**Guilherme Welter Wendt** é doutor em Psicologia pela Universidade de Londres (*Goldsmiths College*), bacharel em Psicologia (PUCRS), mestre em Psicologia Clínica (UNISINOS) e, atualmente, bolsista de pós-doutorado Jr. do CNPq no Instituto de Psicologia da UFRGS.

**Irani I. de Lima Argimon** é psicóloga, doutora em Psicologia, mestre em Educação, especialista em Avaliação Psicológica (PUCRS) em Terapia Cognitivo Comportamental (FBTC), prof. titular curso de Psicologia – Ciências da Saúde e do Curso Gerontologia BioMédica – Medicina e pesquisadora produtividade CNPq.