

O que as escalas de depressão avaliam? Uma discussão sobre similaridade de itens

What do the depression scales evaluate? A discussion of the similarity of items

Makilim Nunes Baptista (orcid.org/0000-0001-6519-254X)¹

Naira Ravanny de Souza Lima²

Resumo

O objetivo do presente artigo foi comparar escalas que avaliam depressão em relação ao conteúdo dos seus itens. Para tanto, as escalas BDI, BDI-II, CES-D e EBADEP-A foram elencadas. Para a avaliação das sobreposições de itens foi utilizado o índice de similaridade de Jaccard. Como resultados, foram identificados 37 sintomas distintos em que a sobreposição dos itens indicou variação de similaridade entre fraca (0,39) a forte (0,73). Discutem-se a aplicabilidade das escalas e o uso mais adequado de acordo com essa característica e as contribuições do estudo para a construção de novos instrumentos. Conclui-se que a similaridade dos itens é importante, pois os instrumentos avaliam o mesmo construto, contanto, dissimilaridades devem ser consideradas no momento da escolha do instrumento a partir do contexto de aplicação e do objetivo da avaliação.

Palavras-chave: 1. Sintomas depressivos. 2. BDI. 3. BDI-II. 4. CES-D. 5. EBADEP-A.

Abstract

This article aimed to compare scales that evaluate depression regarding the content of their items. The BDI, BDI-II, CES-D, and EBADEP-A scales were used. Jaccard's similarity index was used for the evaluation of item overlays. A total of 37 distinct symptoms were identified in the results and the overlapping of the items indicated a similarity variation between weak (0.39) and strong (0.73). It discusses the applicability of the scales and the most appropriate use according to this characteristic and the contributions of the study to the construction of new instruments. It concludes that the similarity of the items is important, since the instruments evaluate the same construct, however, dissimilarities must be considered when choosing the instrument from the context of application and objective of the evaluation.

Keywords: 1. Depressive symptoms. 2. BDI. 3. BDI-II. 4. CES-D. 5. EBADEP-A.

A depressão é um transtorno mental cujos sintomas principais estão relacionados ao humor depressivo (triste, irritado, sentimento de vazio) e perda de interesse ou de prazer (anedonia). Sintomas adicionais estão relacionados às alterações cognitivas, comportamentais e neurovegetativas. O conjunto de sintomas deve durar no mínimo duas semanas e geralmente perdura por vários meses sem alteração significativa de humor. O Manual Diagnóstico e Estatístico de Transtornos Mentais - DSM-5 (American Psychiatric Association, APA, 2014) e a Classificação Internacional de Doenças - CID-11 (Organização Mundial de Saúde, OMS, 2018) geralmente são a base utilizada pelos profissionais para a classificação dos sintomas e do diagnóstico da doença. O indivíduo com depressão experencia alterações no afeto, na cognição e em funções neurovegetativas que devem ocasionar prejuízo no funcionamento social, ocupacional ou em outras áreas importantes da

¹ Universidade São Francisco, Campinas, Brasil. E-mail: makilim01@gmail.com.

² Universidade São Francisco, Campinas, Brasil. E-mail: nairaravannyl@gmail.com.

vida. Além disso, é possível que o transtorno seja recorrente na maioria dos casos (OMS, 2018; APA, 2014) e na pior das consequências, precipitar um suicídio (WHO, 2014).

A CID-11 entrará em vigor em 2022 (OMS, 2018), e dentro das alterações em relação à CID-10 (OMS, 1993) para os critérios diagnósticos de depressão, não houve alteração. A CID-10 (OMS, 1993) faz distinção entre episódio depressivo único e recorrente ao passo que também permite a classificação em três intensidades, a saber, leve, moderada ou grave, em remissão parcial ou completa, podendo haver a presença de sintomas psicóticos ou somáticos. Simultaneamente, o DSM-5 (APA, 2013) apresenta sua classificação categorizando os sintomas de acordo com a quantidade e a intensidade. Segundo o manual, é preciso estar presente cinco do total de nove sintomas, com obrigatoriedade da presença de um ou dois dos sintomas principais.

Tabela 1. Semelhança sintomatológica entre CID e DSM-5

Manual	CID	DSM-5
sintomas	Rebaixamento do humor; Redução da energia e diminuição da atividade; Alteração da capacidade de experimentar o prazer; Perda de interesse; Diminuição da capacidade de concentração, associadas em geral à fadiga importante, mesmo após um esforço mínimo; Problemas do sono; Alteração do apetite; Diminuição da autoestima e da autoconfiança; Ideias de culpabilidade e ou de indignidade; Lentidão psicomotora; Agitação; Perda de peso não induzido; Perda da libido.	Humor deprimido; Diminuição do interesse ou prazer em todas ou quase todas as atividades na maior parte do dia, quase todos os dias; Perda ou ganho significativo de peso sem estar fazendo dieta ou redução ou aumento do apetite; Insônia ou hipersonia; Agitação ou retardo psicomotor; Fadiga ou perda de energia; Sentimentos de inutilidade ou culpa excessiva ou inapropriada (que podem ser delirantes); Capacidade diminuída para pensar ou se concentrar, ou indecisão; Pensamentos recorrentes de morte (não somente medo de morrer), ideação suicida recorrente; Sem um plano específico, uma tentativa de suicídio ou plano específico para cometer suicídio.

Além dos sintomas presentes nos manuais psiquiátricos, existem teóricos que elencam sintomas depressivos complementares que contribuem para um diagnóstico mais preciso e auxiliam no diagnóstico diferencial. Desesperança, pessimismo, angústia, distorções a respeito da própria imagem, passividade e tendência a assumir culpas são alguns exemplos (Baptista, 2012, Solomon, Haaga & Arnow, 2001). Na visão comportamental, Fester, Culbertson e Boren (1977) ressaltam o aumento ou a diminuição da frequência de comportamentos de fuga e esquiva a estímulos como, por exemplo, atividades sociais. Para a Teoria Cognitiva (Beck, Rush, Shaw & Emery, 1997) que fundamenta uma das escalas mais utilizadas mundialmente, o Beck Depression Inventory - BDI, o indivíduo depressivo possui um padrão estável e negativo de funcionamento e falhas de processamento de informação. O sujeito tem uma visão predominantemente negativa de si (autoestima), dos outros (relacionamentos interpessoais) e sobre o futuro (pessimismo), formando assim a Tríade Cognitiva.

Independentemente da teoria, o transtorno apresenta uma variabilidade de perfis, advindos da combinação dos sintomas principais com os sintomas adicionais. Estima-se que possam existir 227 perfis diferentes de pessoas com depressão, combinando variações de presença ou ausência, frequência, intensidade e duração de sintomas. Esse número pode ser ainda maior se forem consideradas nos instrumentos de mensuração respostas tipo Likert (níveis de intensidade), ao invés de dicotômicas (presença/ausência). Assim, para exemplificar, tem-se a situação de dois indivíduos diagnosticados com os mesmos escores em um instrumento que avalia depressão. O primeiro indivíduo pode apresentar tendência suicida, enquanto outro, não; um pode apresentar insônia e o outro, hipersonia; agitação motora ou retardo (Baptista 2018; Dubovisky & Dubovisky, 2004; Fried & Nesse, 2015; Zimmerman, Ellison, Young, Chelminski & Dalrymple, 2015).

Tendo em vista a variabilidade de sintomas depressivos, as escalas de depressão autoaplicáveis nem sempre foram construídas a fim de contemplar todos eles. Isto ocorre porque a construção depende da abordagem que os autores utilizaram. A maioria pauta-se nas considerações do DSM e no modelo Cognitivo de Beck (Wenzerl, Brown & Beck, 2009). Uma forma de teorizar os sintomas é agrupando-os em categorias. Thompson (1989) agrupou os sintomas depressivos em sete categorias: a categoria "Humor" inclui tristeza, perda de interesse e/ou prazer, crises de choro, variação diurna do humor (esta última, às vezes, classificada na categoria de sintomas vegetativos). Essa categoria de sintomas é essencial para o diagnóstico de depressão. A categoria "Vegetativos ou Somáticos" inclui alteração no sono, apetite, peso, libido, fadiga. A categoria "Motores" reúne sintomas de inibição, retardo ou agitação (mais em idosos) e inquietação. A categoria "Sociais" engloba apatia, isolamento, dificuldade para realizar tarefa (incapacitação). A "Cognitivos" agrupa desesperança, desamparo, culpa, ideação suicida, indecisão e perda de insight (não reconhece a doença). Além dessas, o autor apresenta a categoria "Ansiedade", que inclui a

psíquica, somática e fóbica, e a “Irritabilidade”, categoria que reúne hostilidade auto e heterodirigida, também podendo ser considerada na alteração de humor.

No tocante à avaliação, diversos instrumentos foram construídos ao longo dos anos (Santor, Gregus & Welch, 2006). Com o intuito de identificar os instrumentos que avaliam depressão utilizados no contexto brasileiro, apenas três revisões foram encontradas. A primeira foi a de Aros e Yoshida (2009), a partir da base de dados Medline, no período entre 1966 até 2007, utilizando as palavras-chave “depression” e “scale”. Neste estudo, os autores encontraram 49 instrumentos utilizados em pesquisas para todas as faixas etárias (criança, adolescente, adulto e idoso). Os três instrumentos mais citados foram a Hamilton Depression Rating Scale (HAM-D), o Hospital Anxiety and Depression Scale (HADS) e Beck’s Depression Inventory (BDI).

Outra revisão foi a de Ely, Nunes e Carvalho (2014), que fizeram um levantamento em 2013 no Sistema de Avaliação de Testes Psicológicos – SATEPSI, sobre os instrumentos que avaliam depressão com pareceres favoráveis para uso com a população brasileira. Como resultados, 11 instrumentos foram elencados. Dos instrumentos não expressivos, foram encontradas a BDI e a Escala Baptista de Depressão – versão adulta (EBADEP-A).

Em uma busca mais atualizada, Baptista e Borges (2016) revisaram a literatura sobre instrumentos de depressão em crianças/adolescentes e adultos nas bases Index Psi e BVS Psi (SciELO e PePSIC) entre 2005 e 2015. Os autores identificaram 24 instrumentos, sendo 12 avaliando unicamente a depressão, cinco específicos para crianças/adolescentes e os cinco instrumentos mais citados foram BDI, HADS, HAM-D, Children Depression Inventory CDI e EBADEP-A. De acordo com a proposta deste estudo e dos critérios para a escolha das escalas, observa-se que a HADS avalia também ansiedade; HAM-D avalia pacientes com depressão; e CDI envolve população infantil. Então, os testes que avaliam unicamente sintomas depressivos e para a população geral e de idade adulta são BDI e EBADEP-A, sendo estas as escalas escolhidas para este artigo. Também foram as escolhidas a Center of Epidemiologic Studies Depression Scale (CES-D), que foi a conseguinte mais citada nos estudos brasileiros considerando os critérios já apontados, e ainda a BDI-II (Gorenstein, Wang, Argimon & Werlang, 2014) por ser uma versão revisada da BDI, com vistas a abordar também a versão mais recente do instrumento.

Os instrumentos selecionados avaliam unicamente depressão, porém devem-se observar as particularidades do contexto em que estão sendo utilizados. Para exemplificar, tem-se o uso de uma escala em um ambiente hospitalar. Se uma escala foi construída com itens que refletem muitos sintomas somáticos ou vegetativos (alterações de sono por exemplo), esses itens podem ser endossados (reforçados) não necessariamente por um quadro depressivo, mas pelas condições de internação. Da mesma forma que, ao se avaliarem sintomas depressivos em idosos, devem-se considerar sintomas que podem ser apresentados pelas limitações físicas relacionadas à idade do sujeito avaliado (como lentidão psicomotora) e não necessariamente atribuir ao item uma associação direta ao sintoma

depressivo. Calil e Pires (1998) apresentam que o conjunto complexo de sintomas depressivos permitiu que as escalas se diferenciasssem de acordo com a ênfase de alguns sintomas em contraste a outros. Algumas delas priorizam sintomas cognitivos e/ou vegetativos, enquanto outras podem conter mais itens sobre sintomas motores ou de humor, ou ainda, alguns não avaliam, por exemplo, aspectos sociais ou ansiedade. Apesar disto, todas elas são utilizadas como medidas homogêneas, o que não seria adequado.

Verificar a similaridade ou não dos conteúdos de escalas de depressão por meio dos seus itens pode trazer avanços para a aplicação prática, pois auxilia na escolha de quais testes são mais adequados para contextos específicos. Assumindo a similaridade, escalas de classificação só podem ser indicadores equivalentes de gravidade da depressão, na medida

em que os conteúdos dos itens se sobrepõem. É importante que a similaridade aconteça, pois muitos estudos são realizados com escalas diferentes e a não similaridade sugeriria um impedimento acerca da generalização dos resultados e mais um desafio à replicabilidade, pois estariam avaliando de forma diferenciada o mesmo construto, e não se poderia pensar em medidas completamente correspondentes e/ou convergentes (Fried, 2017).

Em uma perspectiva adicional, a não similaridade acarreta maior cuidado quanto à escolha dos instrumentos no contexto de avaliação. Uma vez que há uma variabilidade grande de sintomas depressivos, as escalas podem conter essa diversificação em maior ou menor grau. Dessa forma, o presente artigo tem como objetivo comparar escalas que avaliam depressão, BDI, BDI-II, CES-D e EBADEP-A, em relação ao conteúdo dos itens, avaliando a similaridade entre elas, e propor o uso mais adequado de acordo com as características das escalas. A partir disso, levanta-se a hipótese de que as escalas não são homogêneas em seus conteúdos, baseado na literatura apresentada.

Método

Inventário de Depressão de Beck (BDI; Beck et al., 1961 traduzido por Cunha, 2001)

É um inventário de autorrelato unidimensional que avalia sintomas de depressão em amostras clínicas, não clínicas e psiquiátricas. A escala é unidimensional composta por 21 itens referentes à insatisfação, punição tristeza, sensação de fracasso, pessimismo, culpa autoaversão, autoacusação, retraimento social, indecisão, ideação suicida, irritabilidade, mudança na autoimagem, insônia, dificuldade de trabalhar, perda de apetite e peso, diminuição de libido, fadigabilidade e choro. O índice de confiabilidade foi obtido com seis amostras psiquiátricas e seu valor variou entre 0,79 e 0,90. É respondido em escala Likert de quatro pontos, com variação de zero a três. Possui uma versão revisada, BDI-II (Beck, Steer & Brown, 1996), na qual quatro itens (perda de peso, mudança na autoimagem, preocupações somáticas e dificuldade de trabalhar) foram substituídos por outros quatro novos itens

(agitação, desvalorização, dificuldade de concentração e falta de energia) com o objetivo de identificar sintomas típicos de depressão grave (Gorenstein, et al., 2014).

Center for Epidemiologic Studies CES-D (Radloff, 1977, traduzida por Silveira & Jorge, 2002)

É um instrumento de autorrelato que avalia sintomas depressivos. Possui 20 itens estruturados em quatro fatores oblíquos: afetos positivos (quatro itens), afetos depressivos (sete itens), atividade somática/atividade reduzida (sete itens) e relações interpessoais (dois itens). As respostas são pontuadas em escala Likert de quatro pontos, sendo 1=Raramente ou nunca (menos que 1 dia), 2=Poucas vezes (1-2 dias), 3= Um tempo considerável (3-4 dias), 4=Todo o tempo (5-7 dias). As análises de confiabilidade demonstraram que o instrumento possui boa consistência interna (alfa de Cronbach = 0,85).

Escala Baptista de Depressão - Adulto - EBADEP-A (Baptista, 2012)

É um instrumento de autorrelato que objetiva rastrear sintomas da depressão em amostras psiquiátricas e não psiquiátricas. A escala é unidimensional e composta por 45 itens. Os itens no manual também foram divididos em sete categorias teóricas de sintomas, podendo um mesmo item estar presente em mais de uma categoria. São elas: Social (1) Humor (2) Cognitivos (3) Vegetativos (4) Motores (5) Irritabilidade (6) e Ansiedade (7). O índice de consistência interna do instrumento é adequado, tendo o valor de alfa de Cronbach igual a 0,96. É respondido em formato do tipo Likert de quatro pontos, com variação de zero a três.

Avaliação dos itens por juízes e construção das matrizes

Primeiramente, os itens das escalas foram organizados em uma planilha de Excel a fim de criar duas matrizes de dados. Para tanto, dois pesquisadores, autores do artigo, identificaram, de forma independente, os sintomas contidos nos itens de cada escala (chamado de conteúdo) de acordo com a classificação de 0= conteúdo não estava contido no item, 1= para conteúdo presente parcialmente no item e 2= para conteúdo presente completamente no item. Por exemplo, o item 20 da escala CES-D “Não consegui levar adiante minhas coisas”, no manual recebe classificação como sendo um item de Atividade Somática e para os pesquisadores o item recebeu classificação de 2 para o conteúdo “Incompetência” por julgarem o conteúdo “Incompetência” completamente inserido no item, 1 para “Falta de energia” e 1 para “Perda de interesse” por julgarem esses conteúdos presentes parcialmente no item. Essa reclassificação dos itens pelos seus conteúdos foi feita para todos os itens de todas as escalas. Verifica-se nesse exemplo que um mesmo item poderia receber mais de uma classificação caso houvesse mais de um conteúdo contido nele.

Para a realização da tarefa, os itens com escrita e conteúdo muito similares foram considerados como iguais. Foram eles, 5 e 8 da BDI (“sinto-me culpado o tempo todo” e “culpo-me por todas as coisas ruins que acontecem”, respectivamente) e 1 e 6 da EBADEP-A (“tenho sentido vontade de chorar” e “tenho chorado”, respectivamente). Além disso, para os itens invertidos considerou-se o conteúdo contrário ao item, por exemplo, “sinto-me feliz” foi computado como humor triste ou deprimido. Cabe ressaltar que a classificação do item descrita no manual serviu de base para as análises de identificação dos conteúdos, a fim de não se desviar dessa nomenclatura primária. A partir dessa classificação em conteúdo ausente (=0), parcial (=1) ou totalmente presente no item (=2), verificou-se a concordância da classificação dos pesquisadores e, quando não foram iguais, os pesquisadores avaliaram novamente os itens, dessa vez de forma conjunta chegando ao consenso de 100% quanto aos conteúdos de sintomas para os itens. Apenas três sintomas não possuíram o consenso na primeira avaliação. A partir desse consenso foi possível transformar a primeira matriz não binária (classificação 0, 1 e 2) em binária (0 e 1).

A segunda matriz computou apenas ausência (aqueles que foram classificados apenas com 0) e presença parcial ou total do conteúdo no item (que foram aqueles que receberam =1 ou = a 2). Dessa forma, a segunda matriz é binária e apresenta apenas ausência do conteúdo (=0) ou presença (=1) para servir de base ao cálculo do índice de Jaccard.

Análise de similaridade

A sobreposição de conteúdo dos itens foi estimada usando o coeficiente de similaridade Jaccard para dados binários que varia de 0 (sem sobreposição entre escalas) para 1 (sobreposição completa). Para o cálculo, as escalas foram distribuídas par a par em função dos sintomas. O Índice Jaccard é calculado por $s / (u_1 + u_2 + s)$, no qual s é o número de itens que compartilham duas escalas e u_1 e u_2 é o número de itens que são exclusivos para cada uma das duas escalas. Para considerar um Jaccard forte ou fraco foi utilizado a proposição de Evans (1996) para a coeficiente de correlação: muito fraco ou nulo (entre 0e 0,19), fraco (0,20 a 0,39), moderado (0,40 a 0,59), forte (0,60 a 0,79) e muito forte (0,80 a 1). As análises foram realizadas utilizando uma planilha do Excel a partir dos resultados da análise de conteúdo.

Resultados

Primeiramente, foi verificada a porcentagem de itens para a classificação categórica de Thompson (1989) com o intuito de analisar quais categorias são contempladas pelos instrumentos. Os resultados podem ser vistos na Tabela 2.

Tabela 2. Porcentagem de contribuição das várias categorias de sintomas ao escore máximo possível de cada escala de avaliação de depressão de acordo com a classificação de Thompson (1989)

Categoria	BDI (%)	BDI-II (%)	CES-D (%)	EBADEP-A (%)
Humor	9	9	45	24,5
Vegetativos	29	27	25	15
Motores	0	2	15	4
Sociais	5	5	15	28
Cognitivos	52	52	0	22,5
Ansiedade	0	0	0	2
Irritabilidade	5	5	0	4
Total de itens	21	21	20	45

Observa-se pelos resultados da Tabela 2, que a maioria das escalas contém pelo menos quatro das sete categorias apresentadas por Thompson, sugerindo que, para os autores dos instrumentos, essas sejam as categorias centrais. Na etapa de avaliação de juízes, as escalas BDI e EBADEP-A ficaram com 20 e 44 itens para serem avaliados seguindo o critério de que ambas tiveram alguns itens com escrita similares, a saber BDI itens 5 e 8 e EBADEP-A itens 1 e 6. Para as outras escalas, itens similares para agrupamento não foram identificados, então foi considerada a quantidade inicial de itens, a saber, 21 para BDI-II e 20 para CES-D. Assim, o somatório de todas as escalas resultou em 105 itens para as análises independentes pelos pesquisadores. A partir da construção da primeira matriz (não binária), os pesquisadores entraram em consenso de 37 sintomas distintos que foram identificados no conteúdo dos itens, que podem ser observados na Tabela 3.

Tabela 3. Sintomas identificados a partir da avaliação do conteúdo dos itens.

Categoria (Thompson, 1989)	Conteúdos/Sintomas identificados
Humor	humor triste ou deprimido perda de interesse perda de prazer choro solidão tédio
Cognitivos	inutilidade/incompetência dificuldade de concentração enfrentamento de problemas desesperança; pessimismo

	inadequação
	autocrítica
	autoestima
	indecisão
	ideação suicida
	culpa
	fracasso passado
	medo, pânico
	punição (externo)
Motores	retardo psicomotor
	agitação psicomotora
	perda de energia
	esforço
Irritabilidade	irritabilidade
Sociais	dependência de terceiros
	incômodo
	isolamento
	falar menos
	dificuldade nas relações interpessoais
Vegetativos	insônia
	hipersonia
	apetite diminuído
	apetite aumentado
	libido diminuída
	peso diminuído
Ansiedade	preocupação somática
	hipocondria

Tais sintomas poderiam estar presentes nos itens de duas formas: 1) conteúdo presente totalmente no item e 2) conteúdo presente parcialmente no item. Dessa forma, verificou-se a proporção dos 37 sintomas nas escalas. A concordância total (100%) entre os avaliadores foi utilizada como critério de identificação dos sintomas nos itens, e quando houve discordância, os avaliadores entraram em consenso posteriormente. Os resultados podem ser verificados na Tabela 4.

Tabela 4. Número de sintomas que aparecem combinados nas escalas

Quantidade de Escalas	Quantidade de Sintomas	%
01 (qualquer uma das escalas)	10	27

02 (duas escalas)	07	19
03 (três escalas)	09	24
04 (quatro escalas)	11	30

A Tabela 4 mostra quantos sintomas se agrupam em duas, três ou quatro escalas e quantos aparecem em apenas uma escala. Verifica-se que 30% de sintomas estão presentes no subconjunto das 4 escalas. São eles: inutilidade/incompetência, desesperança/pessimismo, autocrítica, autoestima, fracasso passado, choro, humor triste ou deprimido, perda de energia, insônia, perda de interesse e apetite diminuído. Observa-se que também que 27% dos sintomas estão presentes em uma escala. Os sintomas são citados como: tédio (BDI), preocupação somática (BDI), apetite aumentado (BDI-II), medo/pânico (CES-D), incômodo (CES-D), falar menos que o habitual (CES-D), dificuldade nas relações interpessoais (CES-D), inadequação (EBADEP-A), dependência de terceiros (EBADEP-A), retardo psicomotor (EBADEP-A). Na Tabela 5 estão dispostas as quantidades de itens que cada escala apresentou separadamente, a quantidade de sintomas idiossincráticos, ou seja, o sintoma que foi exclusivo para uma escala e a proporção dos itens da escala em relação ao total de itens elencados.

Tabela 5. Número de sintomas que aparecem combinados nas escalas

Escala	Quantidade ajustada de Sintomas	Sintomas idiossincráticos (%)	Sintomas comuns (%)
BDI	25	05	68
BDI-II	24	03	65
CES-D	18	01	49
EBADEP-A	28	08	76

De acordo com os resultados apresentados na Tabela 4, a EBADEP-A apresentou maior variabilidade de sintomas (n=28) em relação às demais escalas. Em seguida foi a BDI (n=25), BDI-II (n=24), e por fim, a CES-D (n=18). A mesma ordem foi observada para a porcentagem de sintomas comuns em relação ao conjunto total de sintomas, na qual a EBADEP-A deteve 76% dos sintomas comuns. A porcentagem de sintomas idiossincráticos foi pouco expressiva. Nesse quesito, a EBADEP-A também apresentou a maior parte de sintomas que foram identificados apenas para uma escala. Considerando os sintomas diferentes para cada combinação de escalas (par a par) e os sintomas comuns, foi calculado o valor do índice de Jaccard para cada combinação. Tais valores são apresentados na Tabela 6.

Tabela 6. Valores de Jaccard par a par

	BDI	BDI-II	CES-D	EBADEP-A
--	-----	--------	-------	----------

BDI	1			
BDI-II	0,69	1		
CES-D	0,39	0,55	1	
EBADEP-A	0,66	0,73	0,39	1

Considerando os valores de Jaccard de acordo com a proposição de Evans (1996) para os coeficientes de correlação, observa-se, a partir dos resultados da Tabela 6, que a variação de sobreposição das escalas foi entre fraca (0,39) a forte (0,73). As que mais se assemelharam foram a BDI-II e EBADEP-A (0,73), enquanto as menos semelhantes foram CESD-D e BDI (0,39) e CES-D e EBADEP-A (0,39).

Discussão

O objetivo do presente artigo foi comparar escalas que avaliam depressão em relação ao conteúdo dos itens, avaliando a similaridade entre elas. Os resultados demonstram que todas as escalas contemplaram ao menos quatro das sete categorias elencadas por Thompson (1989). Isso denota o cuidado dos autores dos testes em abranger não somente os sintomas principais do transtorno depressivo, como também sintomas complementares.

Conforme as análises qualitativas, foram identificados 37 conteúdos/sintomas a partir dos itens das escalas. Observa-se então que a construção das escalas se pauta não somente nos critérios dos manuais psiquiátricos, mas também, em outras teorias, como a teoria cognitiva. Nas análises quantitativas, verificou-se que 30% de sintomas estão presentes nas 4 escalas abordadas no estudo, ao passo que 27% estavam presentes em uma escala e que se assume a similaridade baixa entre as escalas. As vantagens em termos de similaridade se dão pela abrangência de escolha sobre qual escala o profissional poderá adotar em seu campo de trabalho. Além de estudos adequados de evidência de validade, é importante atentar ao conteúdo dos itens. Caso haja similaridade moderada a alta, as escalas podem ser utilizadas como substitutas. De acordo com Snaith (1993) e Fried (2017), a prática rotineira de usar escalas como medições permutáveis da gravidade da depressão, sem observar a similaridade, é problemática e pode representar ameaça à generalização e replicabilidade da depressão como pesquisa, pois descarta seu aspecto de transtorno com uma variabilidade elevada de manifestações.

A diferença das escalas dá-se pela base de construção das mesmas. A maioria é respaldada pelas versões do DSM e a Teoria Cognitiva de Beck. Entretanto, o DSM tem atualizações e nem sempre as escalas acompanham as atualizações concomitantemente. As análises demonstraram que as escalas com o conteúdo dos itens mais sobrepostos entre si foram a BDI-II e EBADEP-A. O que significa afirmar que a permuta entre essas escalas é melhor do que quaisquer outras. Nota-se que esse resultado inferior à similaridade entre BDI

com BDI-II. Isto é explicado pelo fato de a BDI-II possuir mais itens que descrevem mais conteúdos (sintomas), tal qual a EBADEP-A, do que a BDI.

Outro resultado do estudo foi de que a escala mais “completa” em termos de sintomatologia foi a que apresentava maior quantidade de itens coincidentemente, EBADEP-A. Fried (2016) apresenta que são importantes escalas que sejam abrangentes em termos de quantidade de sintomas, não necessariamente de itens, pois essa característica auxilia no acompanhamento de pacientes depressivos a longo prazo, uma vez que, com vários sintomas, pode-se avaliar melhor a responsividade ao tratamento. Dessa forma, para um psicólogo clínico, sugere-se o uso dessa escala para observância da flutuação de sintomas durante o acompanhamento terapêutico.

Pode-se afirmar que o objetivo proposto foi alcançado. Em relação às limitações, ressalta-se o fato de a análise de categorização ter sido realizada de forma qualitativa, por conta disso, com finalidade de minimizar os possíveis vieses, a mesma foi realizada por dois pesquisadores da área, de forma independente, e os dados trabalhados a partir da concordância dos mesmos. Como proposta para novos estudos, incentiva-se pesquisas abordando instrumentos de depressão para outras faixas etárias, uma vez que este estudo adotou instrumentos voltados para população adulta. Os achados deste estudo refletem a consideração sobre a construção de instrumentos de forma a contemplar os sintomas depressivos ou, caso não seja o objetivo da construção abordar todos eles, que seja então apresentada uma justificativa plausível sobre por que abordar alguns sintomas em detrimento a outros e que esse aspecto seja indicado com mais veemência.

A vantagem de um instrumento com uma gama ampla de sintomatologia permite que, em um exemplo, dois casos tenham a mesma pontuação na escala, sendo que em um o sujeito apresente mais sintomas cognitivos do que outro, com sintomas mais vegetativos, e isso implica diretamente na abordagem clínica e medicamentosa. Uma escala mais restrita em quantidade de itens tem a vantagem de ser mais rápida para casos de pesquisa e triagem, porém pode gerar falsos negativos se não houver atenção para contemplar a amplitude dos sintomas. A partir do demonstrado neste estudo, reforça-se que a avaliação de sintomatologia depressiva requer um olhar amplo sobre os próprios itens da escala, além de aspectos conhecidos como estudos de evidência de validade e tabelas normativas. Aposto a esta premissa, o uso de instrumentos na avaliação psicológica requer cuidados para se pensar também em diagnósticos diferenciais, pois se concorda que os sintomas presentes em transtornos depressivos também podem estar presentes em outras condições clínicas como, por exemplo, transtornos de ansiedade ou síndrome de Burnout.

Financiamento

Esta pesquisa foi financiada pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPQ) e pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes).

Referências

- American Psychiatric Association, APA. (2014). *Manual Diagnóstico e Estatístico dos Transtornos Mentais – DSM-5*. Porto Alegre: Artmed.
- Aros, M. S., & Yoshida, E. M. P. (2009). Estudos da depressão: instrumentos de avaliação e gênero. *Boletim de Psicologia*, (59)130, 61-76. Recuperado de http://pepsic.bvsalud.org/scielo.php?script=sci_arttext&pid=S0006-59432009000100006&lng=pt&tlng=pt.
- Baptista, M. N. (2018). Avaliando "depressões": dos critérios diagnósticos às escalas psicométricas. *Avaliação Psicológica*, 17(3), 301-310. DOI: [10.15689/ap.2018.1703.14265.03](https://doi.org/10.15689/ap.2018.1703.14265.03).
- Baptista, M. N. (2012). *Manual da Escala Baptista de Depressão (versão adulto) – EBADEP-A*. São Paulo: Vetor editora.
- Baptista, M. N. & Borges, L. (2016). Revisão integrativa de instrumentos de depressão em crianças/adolescentes e adultos na população brasileira. *Revista Avaliação Psicológica*, (15)n. esp, 19-32 Recuperado de: http://pepsic.bvsalud.org/scielo.php?script=sci_arttext&pid=S1677-04712016000300004&lng=pt&tlng=pt.
- Baptista, M. N., Souza, M. S., & Alves, G. A. S. (2008). Evidências de validade entre a Escala de Depressão (EDEP), o BDI e o Inventário de Percepção de Suporte Familiar (IPSF). *Psico-USF*, (13)2, 211-220. <https://doi.org/10.1590/s1413-82712008000200008>
- Beck, A. T.; Rush, A. J.; Shaw, B. F. & Emery, G. (1997). *Terapia cognitiva da depressão*. Porto Alegre: Artes Médicas.
- Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *BDI: IIBDI: II Beck depression inventory Manual*. (2ed). San Antonio: The Psychological Corporation.
- Beck, A. T., Ward, C. D., Mendelson, M., Mock, J. & Erbaugh, J. (1961). An inventory for measuring depression. *Archives of general psychiatry* (4) 461-571.
- Calil, M. H & Pires, M. L. N. (1998). Aspectos gerais das escalas de avaliação de depressão. *Revista de Psiquiatria Clínica* (25) 5, p. 240-244. Impresso.
- Cunha, J. A. (2001). *Manual da Versão em Português das Escalas Beck*. São Paulo: Casa do Psicólogo.
- Dubovsky, S. L., & Dubovsky, A. N. (2004). Diagnóstico dos Transtornos de Humor. Em: S. L. Dubovsky, & A. N. Dubovsky. *Transtornos do Humor*. (p.27-103). Porto Alegre: Artes Médicas.

- 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.
- Evans, J. D., 1996. *Straightforward Statistics for the Behavioral Sciences*. Cole Publishing, Pacific Grove, CA.
- Fester, C. B., Culbertson, S., & Perrot Boren, M. C. (1977). *Princípios do comportamento*. São Paulo: Hucitec.
- Fried, E. I. (2017). The 52 symptoms of major depression: Lack of content overlap among seven common depression scales. *Journal of Affective Disorders*, (208), 191-197. <https://doi.org/10.1016/j.jad.2016.10.019>
- Fried, E. I. (2016). Are more responsive depression scales really superior depression scales?. *Journal of Clinical Epidemiology* (77), 4-6. <http://dx.doi.org/10.1016/j.jclinepi.2016.05.004>
- Fried, E. I., & Nesse, R. M. (2015). Depression is not a consistent syndrome: An investigation of unique symptom patterns in the STAR*D study. *Journal of Affective Disorders*, 172, 96-102. <https://doi.org/10.1016/j.jad.2014.10.010>
- Gorenstein, C., Wang, Y. P., Argimon, I. I. L. & Werlang, G. (2014). *Inventário de Depressão Beck - BDI-II*. Manual. São Paulo: Casa do Psicólogo.
- Organização Mundial de Saúde, OMS (2018). *Classificação dos Transtornos Mentais e de Comportamento da CID*. Recuperado de <https://www.paho.org/bra/>
- Organização Mundial de Saúde, OMS. (1993). *Classificação dos Transtornos Mentais e de Comportamento da CID-10*. Porto Alegre: Artmed.
- Radloff L. S. (1977). The CES-D scale: a self-report depression scale for research in the general population. *Applied Psychological Measurement*. (1), 385-401. <https://doi.org/10.1177/014662167700100306>
- Santor, D. A., Gregus, M., & Welch, A. (2006). Eight decades of measurement in depression. *Measurement: Interdisciplinary Research and Perspectives*, (4)3, 135-155, https://doi.org/10.1207/s15366359mea0403_1
- Silveira, D. X. & Jorge, M. R. (2002). Propriedades psicométricas da escala de Rastreamento Populacional para Depressão CES-D em populações clínicas e não-clínicas de adolescentes e adultos jovens In: Gorenstein, C., Andrade L. H. S. G. & Zuarde, A.W. *Escalas de avaliação clínica em psiquiatria e farmacologia*. 125-135. São Paulo: Lemos Editorial.
- Snaith, P. (1993). What do depression rating scales measure?. *Journal of Psychiatry*. 163, 293-298.

- Solomon, A., Haaga, D. A. F., & Arnow, B. A. (2001). Is Clinical Depression Distinct from Subthreshold Depressive Symptoms? A Review of the Continuity Issue in Depression Research. *The Journal of Nervous and Mental Disease*, (189)8, 498-506. <https://doi.org/10.1097/00005053-200108000-00002>
- Thompson, C. (1989). Affective disorders. In *The instruments of psychiatric research*. Thompson, C. London: John Wiley & Sons Ltd. 87-126.
- Wenzel, A., Brown, G., & Beck, A. (2009). *Cognitive for Suicida Therapy Patients Scientific and Clinical Applications*. American Psychological Association: Washington, DC. Recuperado de https://www.mapadaviolencia.org.br/pdf2014/Mapa2014_JovensBrasil.pdf
- World Health Organization (2014). *Preventing suicide: a global imperative*. Geneva, Switzerland: Recuperado de <https://apps.who.int/iris/handle/10665/131056> .
- Zimmerman, M., Ellison, W., Young, D., Chelminski, I., & Dalrymple, K. (2015). How many different ways do patients meet the diagnostic criteria for major depressive disorder? *Comprehensive Psychiatry*, (56), 29-34. <https://doi.org/10.1016/j.comppsy.2014.09.007>

Recebido em: 7/04/2019

Aprovado em: 13/08/2019