Journal of Affective Disorders 283 (2021) 317–324

Contents lists available at ScienceDirect

Journal of Affective Disorders

journal homepage: www.elsevier.com/locate/jad

Research paper

Do self-criticism and somatic symptoms play a key role in chronic depression? Exploring the factor structure of Beck depression inventory-II in a sample of chronically depressed inpatients.

Andreas Høstmælingen [a,*], Pål Ulvenes [b], Helene Amundsen Nissen-Lie [c], Mikkel Eielsen [d], Bruce E. Wampold [e]

[a] Department of Psychology, University of Oslo
[b] Department of Psychology, University of Oslo and Modum Bad Research Institute
[c] Department of Psychology, University of Oslo
[d] Department of Medicine, University of Oslo and Modum Bad Research Institute
[e] University of Wisconsin-Madison and Modum Bad Research Institute

ARTICLE INFO

Keywords:
Beck Depression Inventory-II
Factor analysis
ESEM, Chronic depression
Inpatient, treatment

ABSTRACT

Background: The factor structure of depression differs for different sub-samples. The purpose of this study was to explore the factor structure of Beck Depression Inventory-II in patients with chronic depression presenting for inpatient treatment. Methods: Using exploratory structural equation modeling (ESEM), we explored whether a two-factor solution or a bifactor solution provided best model fit for a sample of 377 patients. For the best fitting model stability was assessed with tests for invariance across primary diagnosis (persistent depressive disorder v. recurrent major depressive disorder), and presence of comorbidity. Results: A bifactor solution with one general factor and two specific factors provided best model fit. Invariance analyses provided support for measurement invariance and stability of the factor solution. Limitations: The naturalistic study design implies some uncertainty regarding possible systematic differences between the patients on demographic and clinical characteristics. Conclusion: The factor structure in our sample was best explained by a general depression factor, one specific factor pertaining to self-criticism, and one consisting of the somatic items fatigue, disturbance of sleep, and appetite. Clinicians could benefit from paying special attention to the subfactors identified, as these findings may have implications for treatment choice for patients with chronic depression.

Introduction

Subsamples of depressed patients seem to vary in symptom profiles reflecting possible subtypes of depression that in turn might respond differently to treatment (Huang & Chen, 2015; Shafer, 2006). It is therefore important to extend the body of literature describing the underlying structure of depression in different patient subsamples. A common approach to understanding the disorder involves examining its latent structures via factor analysis of symptom measures. Descriptions of depression as consisting of depressed affects, self-deprecating cognitions and somatic symptoms can be traced back to Hippocrates (Spielberger, Ritterband, Reheiser, & Brunner, 2003). However, the factor analytic literature on Beck Depression Inventory-II (BDI-II), the most commonly used depression instrument (Lemmens, Müller, Arntz, & Huibers, 2016), rarely identifies three distinct factors in clinical psychiatric samples. The original study by Beck, Steer, and Brown (1996) identified a two-factor solution consisting of a 9-item cognitive factor and a 12-item somatic-affective factor. Reviews of subsequent studies show three-factor solutions have been identified in samples of substance abusers, post-partum women, students, chronic pain patients, patients with intellectual disabilities and other medical samples, but two-factor solutions are typically identified in clinical psychiatric and depressed samples (Huang & Chen, 2015; Wang & Gorenstein, 2013).

The two-factor solutions are variations of cognitive, somatic and affective elements making up the factors, but different item compositions interfere with straightforward interpretation (Vanheule, Desmet, Groenvynck, Rosseel, & Fontaine, 2008). While some items are consistent indicators of the cognitive dimension, and some consistently define the somatic dimension, other items variably load on one factor or the other to produce either a Cognitive-affective factor or a

* Corresponding author at: Forskningsveien 3a, 0373 Oslo.

https://doi.org/10.1016/j.jad.2021.01.066
Received 28 September 2020; Received in revised form 6 January 2021; Accepted 30 January 2021
Available online 2 February 2021

Somatic-affective factor (Ward, 2006). Using the nomenclature of Beck et al. (1996), these shifting items could be classified as "affective" (Ward, 2006). One reason for the instability of the affective items (such as "sadness", "agitation", "irritability", and "loss of pleasure") across samples might be that they are ambiguous in nature with the ability to add salience to both thought content and non-verbal bodily sensations. Different negative thoughts (e.g., "I am disappointed in myself", "I feel guilty") can add meaning to the circumstances under which negative affect is experienced, and thus become depressive thoughts (Spielberger et al., 2003). Conversely, symptoms such as tiredness/fatigue, changes in sleeping pattern or changes in appetite, may shift from neutral to negative experiences when they appear in conjunction with negative affect. Thus, rather than functioning as a separate factor in depression, affective symptoms (i.e., negative feelings) may add salience to thought content or bodily sensations in different subsamples, making up either cognitive-affective or somatic-affective factors.

One problem with first-order factor solutions is that they fail to represent multidimensionality that occurs when indicators are associated with more than one construct (Morin, Arens, & Marsh, 2016). This is often the case for items in scales measuring psychological constructs (Morin et al., 2016). For example, in an intelligence test some items might be expected to be associated with a sub-domain (e.g., verbal intelligence) as well as to a hierarchically superior construct (e.g., global intelligence). This raises the question whether some depression symptoms, such as affective symptoms, are part of a global construct while other symptoms constitute specific sub-factors in different subsamples of depressed patients. A bifactor model directly tests whether a global construct (a 'g factor') exists as a unitary dimension underlying the response to all items and coexists with specific factors explaining the residual variance not explained by the g factor (Morin et al., 2016). Some studies have reported that bifactor solutions of the BDI-II provide better fit compared to previously identified two-factor solutions in psychiatric outpatients (Brouwer, Meijer, & Zevalkink, 2013), depressed outpatients (Quilty, Zhang, & Bagby, 2010), and psychiatric inpatients (Subica et al., 2014). Also, re-analyses of data from previous studies finding support for two-factor solutions, have found improved model fit when testing a bifactor model (i.e., with one higher-order general factor and two lower order factors; Ward, 2006). Findings supporting bifactor models for BDI-II, corroborate the theory that BDI-II assesses generalized distress along with more specific features of depression (Subica et al., 2014).

Chronic depression (CD) is not a formal diagnosis in current diagnostical classification manuals, but the term is frequently used to describe patients who experience a repeated pattern of recurrent episodes as well as persistence of symptoms (e.g., Jobst et al., 2016; Köhler, Chrysanthou, Guhn, & Sterzer, 2019). It is likely that the pathogenesis of single episode depression is different from that of recurrent and persistent depression, which is characterized by long-term declines in functioning and cognition (Belmaker & Agam, 2008). Also, similar risk factors (e.g., initial depressive and comorbid symptom severity, failure to seek treatment at baseline), predict both persistence (i.e., continuity of symptoms over at least two years) and recurrence of depressive episodes (Hoertel et al., 2017; ten Have et al., 2018). Thus, patients diagnosed with persistent depressive disorder (PDD) and recurrent major depressive disorder (rMDD) are often included in studies exploring chronic forms of depression (Barnhofer et al., 2009; Bockting et al., 2005; DeRubeis et al., 2020; Hollon et al., 2014; Humer et al., 2020; Ma & Teasdale, 2004). On the other hand, PDD and rMDD are clearly separated as two distinct disorders in current diagnostic manuals (American Psychiatric Association, 2013), and there is little agreement on the number and nature of depression subtypes (Fried & Nesse, 2015). Whether clustering of PDD and rMDD is a valid way of conceptualizing chronicity of depression thus remains an open question. Examining whether patients with these diagnoses share similar symptom structures may contribute to the debate on how best to conceptualize chronicity of depression.

For patients diagnosed with depression, prevalence estimates indicate 93.5 percent of them experience at least one other comorbid physical or mental disorder, and patients' evaluations of their own burden of disease are dramatically improved when adjusting for comorbidity (Gadermann, Alonso, Vilgaut, Zaslavsky, & Kessler, 2012). This suggests condition specific severity varies significantly depending on the presence or absence of comorbidity (Moussavi et al., 2007). Also, failure to identify underlying causes of mental disorders suggests they could be understood as clusters of mutually re-enforcing symptoms (Borsboom & Cramer, 2013; Kendler, Zachar, & Craver, 2011). Hence, the presence of comorbid conditions in conjunction with depression may constitute large clusters of re-enforcing symptoms affecting overall symptom severity, functioning and perceived wellbeing, raising the question whether depressed patients with comorbid diagnoses may have different factor structures than patients without comorbidity.

To summarize, it is important to extend the body of literature describing the factor structure of commonly used depression screening instruments for different patient subsamples. Specifically, there is a need to explore the underlying constructs for patients with chronic depression, and whether symptom structure differs between patients with PDD v. rMDD and comorbidity v. no comorbidity.

Previous studies exploring BDI-II have regularly been conducted using variations of exploratory (EFA) and confirmatory (CFA) factor analysis (Huang & Chen, 2015; Wang & Gorenstein, 2013). However, EFA and CFA have methodological limitations (Asparouhov & Muthén, 2009; Marsh, Morin, Parker, & Kaur, 2014). Cross-loadings are traditionally constrained to be zero in CFA but are freely estimated in EFA, so CFA structures are more restrictive than EFA structures. Because of this, in many instances item-level CFAs fail to provide clear support for instruments that have been well established in EFA research (Marsh et al., 2014). Also, the independent cluster model inherent in CFA (ICM-CFA) in which items are required to load on only one factor, could be too restrictive for many multidimensional constructs (Morin et al., 2016). Exploratory structural equation modeling (ESEM; Asparouhov & Muthén, 2009) allows for integration of EFA within a structural equation modeling (SEM) framework. As in EFA, ESEM allows for items to load freely on all factors but at the same time allowing for methodological advances typically reserved for CFA and SEM, such as goodness of fit statistics and comparison of competing models (Marsh et al., 2014; Morin, Marsh, & Nagengast, 2013). ESEM has provided better fit to data and less differentiated factors than CFA (Morin et al., 2013), and performs better in terms of construct validity of the interpretation of the factor structure (Marsh et al., 2009). However, a first order ESEM model will likely ignore the presence of hierarchically superior constructs, which will end up being expressed through inflated cross-loadings. To fully capture the hierarchical and multidimensional nature of instruments incorporating sources of psychometric multidimensionality bifactor ESEM is a viable option (Morin et al., 2016).

The purpose of this study was to explore the factor structure of BDI-II in a sample of hospitalized inpatients with chronic depression (i.e., primary diagnosis PDD or rMDD), using updated statistical methods. We based our analysis on previous studies indicating BDI-II in adult clinical psychiatric samples is best represented either through one global construct with some symptoms constituting specific sub-dimensions (bifactor model) or a two-factor structure. Hence, we tested whether a two-factor structure or a bifactor structure with one general factor and two lower order factors provided best fit for our data, applying ESEM. We also conducted invariance analyses to examine whether factor structure was stable across primary diagnosis and presence of comorbid disorders. To our knowledge no studies on the factor structure of BDI-II have been made on chronically depressed inpatients using ESEM.

Methods

Study design & treatment context

The factor analysis was conducted as part of a naturalistic study of patients presenting for a 12-week inpatient treatment program for chronic depression at Modum Bad hospital in Vikersund, Norway, comparing outcomes of patients that were taking antidepressant medication (ADM) in addition to undergoing inpatient psychotherapeutic treatment with patients who were not taking ADM. Modum Bad has a nation-wide catchment area and patients were referred from general practitioners or local secondary mental health care units across the country. Patients who had exhausted available local treatment options, typically including both pharmaco- and/or psychotherapy, were assessed for the treatment program during a 4-day assessment stay prior to inclusion in the program. Eligible individuals had PDD or rMDD as primary diagnosis. As the risk of recurrence increases progressively with each new episode (de Jonge et al., 2018), and patients on their third or more episode approaches 100% chance of subsequent recurrence (American Psychiatric Association, 2010), patients with a recurrent depressive episode with at least two previous episodes (i.e., current episode is third or more) were included in the study. Exclusion criteria for the treatment program were 1) psychosis, 2) cluster A and B personality disorder, 3) untreated/unstabilized bipolar disorder, 4) ongoing substance abuse and 5) organic brain disorders. All patients applying for the treatment program were diagnostically assessed with the Mini-International Neuropsychiatric Interview (M.I.N.I; Sheehan et al., 1998) and Structured Clinical Interview for DSM-IV Axis II Personality Disorders (SCID-2; First, Gibbon, Spitzer, Williams, & Benjamin, 1997) the first day of the 4-day assessment stay. The same diagnostic instruments were used throughout the period patients were admitted to the program (from 2012 to 2017). A specialist in clinical psychology or psychiatry conducted the interviews and made initial assessment of primary and secondary diagnoses. Then, each diagnosis was discussed in a team of psychiatrists and psychologists before final diagnostic assessment was recorded.

Participants

Between 2012 and 2017, 1800 patients were referred to the treatment program, of which 1200 were excluded because they had not exhausted local treatment alternatives. These were referred back to alternative local health care alternatives. The remaining 600 patients were assessed for eligibility. Some patients (N=163) were excluded for not meeting criteria for persistent or recurrent depression or met exclusion criteria for the treatment program (see above). Thus, 437 patients received treatment. Because 60 patients did not complete the BDI-II at start of treatment, 377 cases were included in the present analyses (see Fig. 1).

Measures

To assess levels and change of depressive symptoms, patients completed BDI–II at assessment, start of treatment, at termination, and at one-year follow-up. In this study, we used the BDI-II data from start of treatment. The BDI-II consists of 21-items, scored on a Likert scale from 0 to 3 (range 0-63), and has demonstrated high reliability and good concurrent, content, and structural validity for screening depression in outpatient and student samples (Beck et al., 1996). Cronbach's alpha showed good reliability for BDI-II in the current sample (α=0.88).

Statistical procedures

We based our analysis on comprehensive reviews of BDI-II most commonly identifying two-factor solutions in adult clinical psychiatric and depressed samples (Huang & Chen, 2015; Wang & Gorenstein, 2013), and findings suggesting that bifactor solutions provide better fit than previously identified two-factor solutions (Brouwer, Meijer, & Zevalkink, 2013; Quilty, Zhang, & Bagby, 2010; Subica et al., 2014; Ward, 2006). Thus, we conducted two exploratory analyses comparing a two-factor structure to a bifactor structure with one higher order, general factor and two lower order factors. To conduct the analyses, we used exploratory structural equation modeling (ESEM; Asparouhov & Muthén, 2009; Marsh et al., 2014). Thus, we contrasted a first order ESEM model with two factors with a bifactor ESEM specifying one general factor and two sub-factors.

All analyses were conducted in Mplus 8 with maximum likelihood estimator (ML; Muthén & Muthén, 1998-2017). First, an exploratory analysis using ESEM was conducted specifying the extraction of two factors. The factors were correlated under the oblique geomin rotation (Muthén & Muthén, 1998-2017). Secondly, a bifactor exploratory analysis was conducted using ESEM, specifying one general factor and

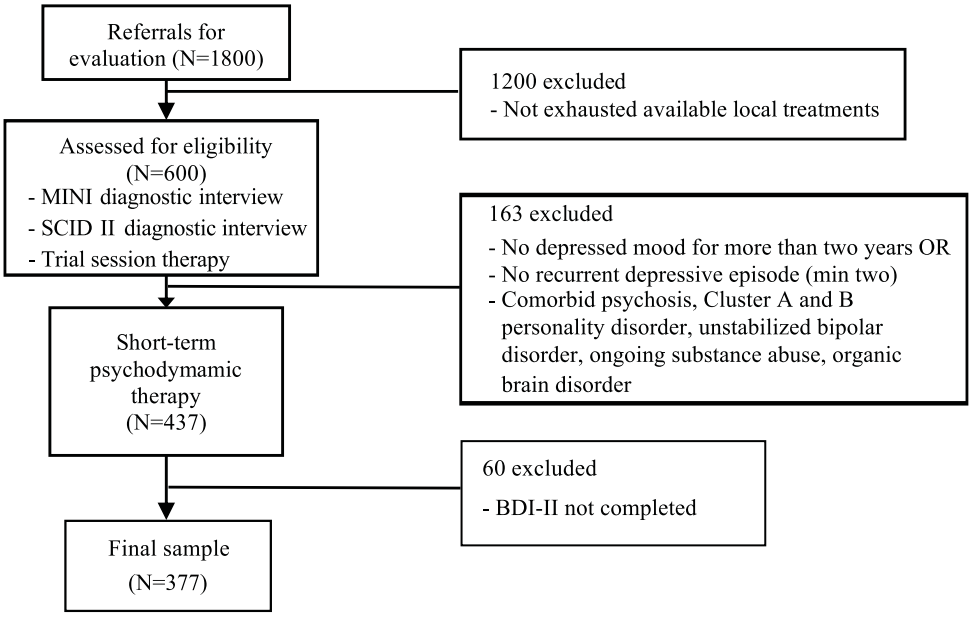

Fig. 1. Study profile.

two specific factors. In bifactor estimation it is assumed that the general and group factors are orthogonal (Reise, 2012). Thus, we specified a bi-geomin orthogonal rotation where the specific factors were uncorrelated. In both models, item loadings were freely estimated, the intercepts and residual variances of the factor indicators were estimated, and the residuals were not correlated. The variances of the factors were fixed at 1 as the default.

For the bifactor model, the independent contributions of general and specific factors to common item variance were determined by calculating the percentage of explained common variance (ECV) for each factor. For each factor the ECV is the sum of the squared standardized factor loadings for that factor divided by the sum of all squared factor loadings for the model (Rodriguez, Reise, & Haviland, 2016). Thus, ECV is the percent of variance explained by each factor.

With a sample size of 377 cases, factor loadings were interpreted as salient when greater than or equal to .30 (Hair, Tatham, Anderson, & Black, 1998). Goodness of fit of the factor model was assessed by means of chi square (χ^2), comparative fit index (CFI), root mean square error of approximation (RMSEA) and standardized root mean square residual (SRMR; Schweizer, 2010). For the CFI cut-offs for acceptable and good model fit we used $\leq$.90 and $\leq$.95, whereas cut-offs for acceptable and good model fit on the RMSEA were set to below .08 or .05 respectively (Marsh et al., 2010). For SRMR, values were expected to stay below 0.10 (Kline, 2005). We used Akaike information criterion (AIC) to compare model fit between the two models.

It is important to establish whether questionnaires measure the same constructs in all subgroups of the population for whom the measure will be used (Brown, 2013). Tests of measurement invariance evaluate the extent to which measurement properties generalize over multiple groups, situations or occasions (Morin, Marsh, & Nagengast, 2013). We tested invariance of the most optimal model across patients with different primary diagnosis (PDD v. rMDD), and comorbidity (comorbid diagnosis present v. not present). First, model fit of the selected model was tested separately in each sub-group (Brown, 2013). Then we sequentially tested configural, weak, strong and strict invariance (Liu et al., 2017; Meredith, 1993; Meredith & Teresi, 2006). Invariance testing was done in MPlus Version 8 following the procedure outlined in Morin et al. (2013, see supplemental materials for Mplus syntax). For analysis of configural invariance factor structures are freely estimated in each group with only the number of factors being the same in both groups. The latent variances are fixed to 1 and the latent means to 0 in both groups to freely estimate all factor loadings and items intercept. Weak invariance tests whether the factor loadings are the same in both groups by fixing the loadings to equality across groups, and fixing factor variance to 1 in a selected reference group while freely estimating it in the other. Strong invariance tests whether intercepts in addition to factor loadings are invariant across groups (i.e., whether individuals with the same score on a latent factor answer the items in a similar way). The intercepts are constrained to equality in both groups, while latent means are constrained to 0 in a selected reference group and freely estimated in the other. Strict invariance requires invariance of item uniqueness (i.e., item-level measurement errors are equivalent across groups) in addition to the invariance of factor loadings and intercepts. This is done by adding equality constraints to item uniqueness in both groups (Morin et al., 2013).

If configural invariance was established, further analysis was conducted to check weak factorial invariance, if weak factorial invariance was established analysis for strong was conducted, and if strong factorial invariance was established, we analyzed for strict factorial invariance. If for any step invariance was not established further analysis was not conducted. If strict invariance is established, this would imply that group differences in means, variances, and covariances of the measured indicators are entirely attributable to group differences in the latent common factors (Millsap, 2011). For purposes of model comparison, tests of the relative fit of models are of greater importance than the absolute level of fit for any one model (Marsh et al., 2009). Differences in comparative fit index (CFI) and root mean square error of approximation (RMSEA) were used as they appear to be equally sensitive to lack of invariance (Chen, 2007). If a difference in CFI is smaller than or equal to .01, this indicates that the hypothesis of invariance is supported (Chen, 2007; Cheung & Rensvold, 2002). For the RMSEA a difference smaller or equal to .015 would support the hypothesis of invariance (Chen, 2007).

Results

The mean BDI-II total score for the sample was 29.47 (*SD = 9.49*) at assessment. The mean age of the patients was 47.5 years (*SD = 10.83*). Years since first episode was 23.5 (*SD = 13.6*), mean 'years since first treatment attempt' was 11.9 (*SD = 9.8*). The primary diagnosis was rMDD for 221 patients (58.6%), and PDD for 156 patients (41.4%). Comorbid psychiatric diagnosis (one or more) was present for 185 patients (49.1%). See Table 1 for additional demographics and clinical characteristics.

Table 2 presents model fit indices for the tested models. Both models provided adequate fit, with the bifactor model achieving the best fit ($\chi 2$ (150) = 250.676, p<.001; RMSEA=0.042, 90% C.I. [0.033, 0.051]; CFI=0.956; SRMR=0.034). Table 3 and Fig. 2 show the factor loadings for the bifactor model. All items except item 16 ("changes in sleeping pattern") loaded saliently (above 0.3) on the general factor. For the first sub-factor, four items loaded saliently (item 5, "guilty feelings"; item 7, "self-dislike"; item 8, "self-criticalness"; item 14, "worthlessness"). As these items all reflect self-devaluating thought content, we labeled this factor "self-criticism". For the second sub-factor, three items loaded saliently (item 16, "changes in sleeping pattern"; item 18, "changes in appetite"; item 20, "tiredness or fatigue"). We labeled this factor

Table 1
Demographic and clinical characteristics.

	Baseline characteristic (N=377)	
	n	%
Sex		
Female	255	67.6
Male	122	32.4
Having children	268	71.1
Marital status		
Single	93	24.7
Relationship	17	4.5
Married or cohabiting	188	49.8
Divorced or widowed	79	21
Education		
No known education	9	2.4
Primary or secondary	24	6.4
High school	87	23.1
Bachelor or higher	257	68.1
Employment status		
Full time work	53	14.1
Part time work	117	31.0
No work	198	52.5
Student	8	2.1
Unknown	1	0.3
Medication at start of treatment		
No medication	171	45.4
Any medication present	206	54.6
Antidepressants	125	33.2
Anxiolytics/hypnotics	38	10.1
Hyperkinetic medication	6	1.6
Mood stabilizers	10	2.7
Antiepileptics	38	10.1
Substance dependency medication	2	0.5
Antipsychotics	34	9.0
Antihistamines	16	4.2
Pain medication	21	5.6
Unknown medication	16	4.2
Primary diagnosis		
Recurrent major depressive disorder (rMDD)	221	58.6
Persistent depressive disorder (PDD)	156	41.4
Comorbid diagnosis	185	49.1

Table 2
Summary of goodness of fit statistics.

	χ2(df)	RMSEA	90% C.I.	CFI	SRMR	AIC	Δ RMSEA	Δ CFI
ESEM	331.376* (169)	0.05	[0.042, 0.058]	0.928	0.04	17673.78		
BI-ESEM	250.676* (150)	0.042	[0.033, 0.051]	0.956	0.034	17631.081		
			Invariance comorbid					
Configural	432.748* (300)	0.048	[0.038, 0.058]	0.942	0.043	17686.445		
Weak	504.949* (354)	0.048	[0.038, 0.057]	0.933	0.057	17650.645	0	0.009
Strong	528.601* (372)	0.047	[0.038, 0.056]	0.931	0.059	17638.298	0.001	0.002
Strict	547.486* (393)	0.046	[0.036, 0.055]	0.932	0.061	17615.183	0.001	-0.001
			Invariance PDD v. recurrent MDD					
Configural	450.177* (300)	0.052	[0.041, 0.061]	0.936	0.043	17708.110		
Weak	523.473* (354)	0.05	[0.041, 0.059]	0.928	0.055	17673.406	0.002	0.008
Strong	542.173* (372)	0.049	[0.040, 0.058]	0.928	0.057	17656.106	0.001	0
Strict	569.460* (393)	0.049	[0.040, 0.057]	0.925	0.06	17641.393	0	0.003

Note. Invariance comorbid: patients with one or more comorbid diagnosis compared to patients with only one diagnosis. Invariance PDD v. recurrent MDD: patients with PDD as primary diagnosis compared to patients with recurrent MDD as primary diagnosis. Estimator is maximum likelihood (ML); ESEM=Exploratory structural equation modeling; BI-ESEM=bifactor ESEM; RMSEA=root mean square error of approximation; C.I.=confidence interval; CFI=comparative fit index; SRMR=standardized root mean square residual; AIC=akaike information criterion; Δ=difference previous model; *p<0.01; ESEM estimated with geomin oblique rotation; Bifactor ESEM estimated with bi-geomin orthogonal rotation.

Table 3
Results from factor analysis of BDI-II.

BDI-II item	Bi-factor ESEM		
	General factor	«Self-criticism»	«Somatic»
1. Sadness	**.624**	-.001	-.187
2. Pessimism	**.557**	.071	-.193
3. Past failure	**.529**	.297	-.100
4. Loss of pleasure	**.660**	-.100	.104
5. Guilty feelings	**.522**	**.443**	.001
6. Punishment feelings	**.438**	.178	-.019
7. Self-dislike	**.556**	**.473**	-.050
8. Self-criticalness	**.521**	**.470**	.107
9. Suicidal thoughts or wishes	**.413**	.154	-.239
10. Crying	**.394**	-.015	.116
11. Agitation	**.389**	.053	.194
12. Loss of interest	**.640**	-.252	.006
13. Indecisivness	**.625**	.019	.229
14. Worthlessness	**.619**	**.300**	-.197
15. Loss of energy	**.643**	-.142	.266
16. Changes in sleeping pattern	.261	.047	**.418**
17. Irritability	**.351**	-.032	.154
18. Changes in appetite	**.358**	.065	**.346**
19. Consentration difficulty	**.631**	-.093	.290
20. Tiredness or fatigue	**.609**	-.129	**.431**
21. Loss of interest in sex	**.320**	-.105	.289

Note. N = 377. The extraction method was exploratory structural equation modeling with maximum likelihood estimator (ML) and orthogonal bi-geomin rotation. Factor loadings above .30 are in bold. Standardized model results.

"somatic". ECV showed 73.4% of the variance of the bifactor model was explained by the general factor, indicating a strong general factor. 13.1% of the variance was explained by the "self-criticism" factor, and 13.5% was explained by "somatic" factor.

Invariance tests for the bifactor model were conducted for presence of comorbid diagnosis v. no comorbid diagnosis, and for primary diagnosis PDD v. recurrent MDD. The bifactor model showed good fit for patients without comorbid diagnosis ($\chi2$ (150) = 208.514, p<.001; RMSEA=0.045, 90% C.I. [0.029, 0.059]; CFI=0.945; SRMR=0.043), and for patients with one or more comorbid diagnosis ($\chi2$ (150) = 224.234, p<.001; RMSEA=0.052, 90% C.I. [0.037, 0.065]; CFI=0.939; SRMR=0.043). For rMDD v. PDD, model fit was good for patients with rMDD as primary diagnosis ($\chi2$ (150) = 221.197, p<.001; RMSEA=0.046, 90% C.I. [0.033, 0.059]; CFI=0.949; SRMR=0.039), and acceptable for patients with PDD as primary diagnosis ($\chi2$ (150) = 228.980, p<.001; RMSEA=0.058, 90% C.I. [0.042, 0.073]; CFI=0.917; SRMR=0.048). For the tests of measurement invariance, the goodness of fit indices suggested good model fit at each stage for all groups (see Table 2). Changes in the goodness of fit indices did not decrease below the limits indicating strong support for measurement invariance and stability of the factor solution (see Table 2).

Discussion

In this study, we explored the factor structure of BDI-II in a sample of inpatients with chronic depression, also testing for invariance between patients with PDD or rMDD as primary diagnosis, and presence of comorbid diagnoses. In our sample we found a high level of symptom severity, a long history of depression, and a long history of treatment attempts. A bifactor model provided best fit, suggesting that psychometric multidimensionality may be present in the BDI-II ratings from our sample. Invariance testing indicated stability of the model across primary diagnosis, indicating the same factor structure for patients with PDD and recurrent MDD. This supports including patients with both diagnoses in studies of chronic depression. The invariance testing also indicated the same factor structure of depression for patients with and without comorbidity. Even though comorbid diagnoses affect symptom severity and perception of burden of disease (Gadermann et al., 2012; Moussavi et al., 2007), our results show that depression remains a stable construct with or without comorbidity present.

Our results further suggest that BDI-II items correspond to one global depression factor, where all items loaded saliently except item 16 ("changes in sleeping pattern"). In addition, some of the items seem to constitute separate sub-dimensions where items revolving around self-critical cognitions (items 5, 7, 8, and 14) load on one specific factor, and somatic items connected to sleep, appetite and fatigue (items 16, 18, and 20) load on another. Also, all of the items typically labelled "affective" loaded on the general factor, but none of the specific factors. Our results indicate depression is mostly explained by a general factor, where affective symptoms are part of the more fundamental (global) construct while cognitive symptoms pertaining to being self-critical and somatic items pertaining to sleep, appetite and fatigue, may play a special role for the current subsample of chronically depressed inpatients.

Among the cognitive aspects of depression, such as helplessness, worrying about the future, ruminating over past problems and self-critical thoughts (Blatt, Quinlan, Chevron, McDonald, & Zuroff, 1982; Pearson, Brewin, Rhodes, & McCarron, 2008), self-criticism may play a particularly important role in chronic/recurrent depression. In an early, study Dent and Teasdale (1988) found thought content, specifically devaluing the self, contributed to chronicity of depression. Also, self-criticism has been linked to severity of depression (Luyten et al., 2007), and higher rates of depressive relapse (Hawley, Zuroff, Brozina, Ho, & Dobson, 2014; Mongrain & Leather, 2006). In addition, less self-criticism and/or greater reduction during inpatient or hospital day

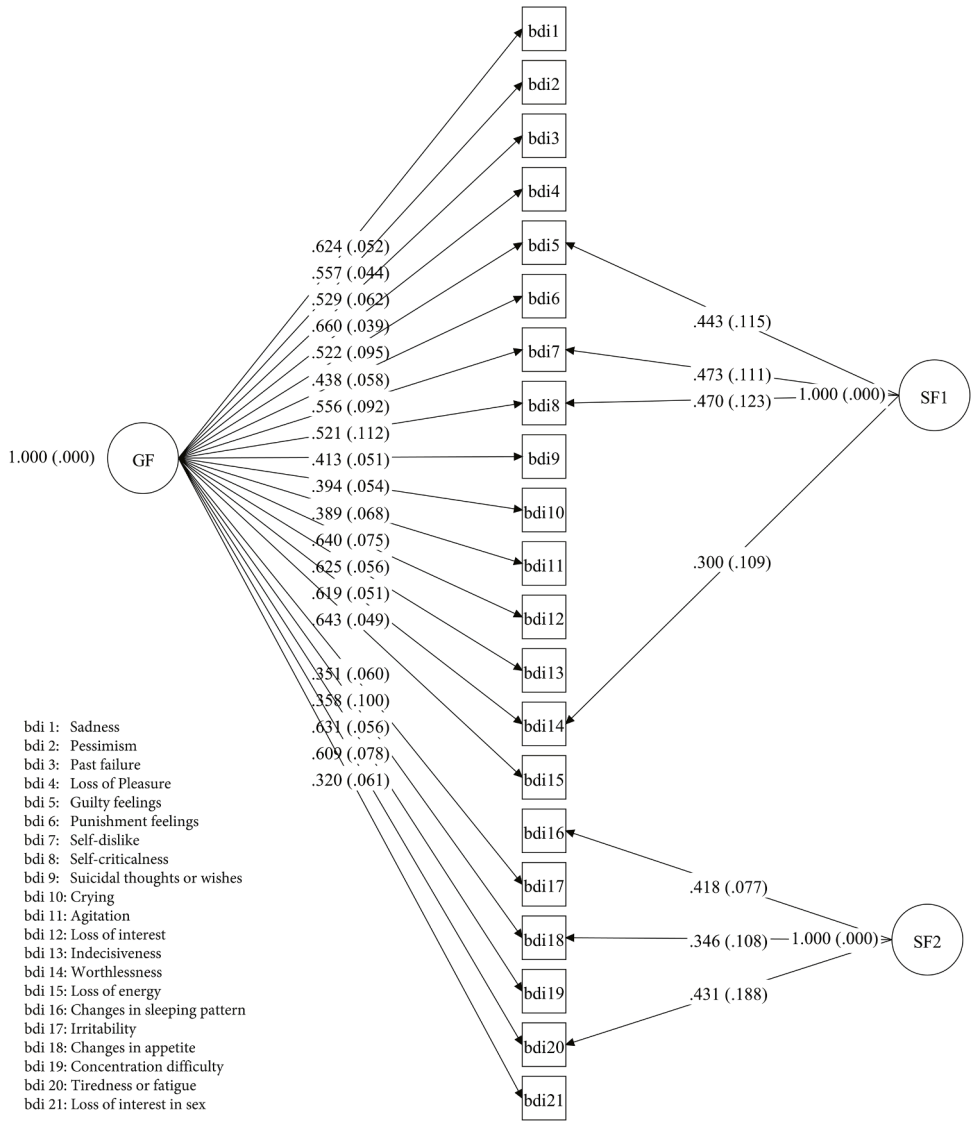

Fig. 2. Bifactor ESEM model.

treatment predicted rapid and sustained improvement after one year for depressed patients (Zeeck et al., 2020). Harsh forms of self-criticism are persistent and difficult to change and may represent a possible specific target for psychotherapeutic treatment (Werner, Tibubos, Rohrmann, & Reiss, 2019).

The three items 16 ("changes in sleeping pattern"), 18 ("changes in appetite") and 20 ("tiredness/fatigue") are the most consistent items regularly loading on a somatic factor (Manian, Schmidt, Bornstein, & Martinez, 2013). One study found that sleep symptoms might be a candidate for one symptom cluster in a "true" symptom structure for depression (Chekroud et al., 2017). From a theoretical perspective, sleep symptoms could directly affect appetite and fatigue forming a set of somatic symptoms that should be specifically addressed in treatment.

Our results could have practical clinical implications. First, the general factor in a bifactor model represents the single source of common variance running through all items in an instrument and can be interpreted as representing the psychological construct the instrument was created to measure (Reise, 2012). According to Beck et al. (1996) the total BDI-II score provides an estimate of the overall severity of depression. Thus, our results indicate that the total BDI-II score is a valid indicator for depression severity in chronically depressed inpatients. However, overall symptom improvement may obscure whether different treatments target different symptoms. Overall, both psychotherapy and antidepressants (ADM) work about equally well for depression (Hollon, 2016), but regardless of treatment type, only 30-40% will achieve remission (Craighead & Dunlop, 2014). In other words, a large number of patients do not respond to either ADM or psychotherapy, and there is a need to identify indicators that predict which patients will respond to different available treatment options (DeRubeis et al., 2014). Specific symptom profiles could serve as indicators for treatment choice (Stewart & Harkness, 2012), and assessing scores on the subfactors self-criticism and somatic items may be useful in guiding treatment choice. For example, positive change due to psychotherapy is most often associated with changes in dysfunctional attitudes, rumination and worry (Lemmens et al., 2016), whereas different ADMs have differential effects on core emotional and sleep symptoms (Chekroud et al., 2017). Also, as depressive symptoms seem to be interconnected in complex networks, improvement or worsening of one type of symptom can causally affect others (Borsboom & Cramer, 2013; Boschloo et al., 2019). For instance, the somatic item sleep disturbance can lead to cognitive impairment (Fried & Nesse, 2015). Thus, targeting specific cognitive or somatic symptoms with treatment options specifically suited for that symptom cluster could lead to overall faster and more stable remission. Further research should explore whether targeting symptom clusters of self-criticism and sleep/appetite/fatigue with different treatment options could provide beneficial outcomes for patients with

chronic/recurrent depression. For instance, if self-criticism is especially salient, therapists may consider exploring psychotherapeutic interventions targeting self-compassion (Neff & Vonk, 2009). If, on the other hand somatic symptoms are particularly salient, one might consider focusing on treatment with antidepressant medication (Chekroud et al., 2017). We believe our results contribute an important clinical nuance in the use and interpretation of BDI-II for chronically depressed patients. Clinicians could benefit from paying special attention to the subfactors identified, as these findings may have implications for the treatment choice for patients with chronic depression.

Limitations

Despite the strengths of this study such as using new statistical approaches and the large sample size of a heterogeneous, naturalistic sample of depressed patients with severe symptomatology, our study has some shortcomings that should be taken into account. First, as this was a naturalistic study, patients were not randomized to treatment conditions leaving uncertainty regarding possible systematic differences between the patients on demographic and clinical characteristics. However, as we did find factorial invariance, any difference between the groups did not affect the main findings of this study. Second, even if our inclusion criteria were liberal, we did exclude those with Cluster B or C personality disorders and those with substance abuse, and so the generalizability of the findings to other samples of depressed patients with these comorbid diagnoses might be compromised. Third, ESEM may have potential limitations, such as not being applicable to complex models unless sample size is sufficiently large (Marsh, et al., 2014). Also, ESEM might confound constructs that need to be kept separate in relation to theory, and ESEM, like EFA, suffers from rotational indeterminacy (i.e., different rotation strategies result in different solution that all fit the data equally well; Marsh, Guo, Dicke, Parker, & Craven, 2020). Lastly, even though we replicated findings from prior studies on BDI-II in a sample of chronic depression, there could be differences in phenomenology of depression that the BDI-II does not capture that may distinguish this group from other groups of depressed patients, and potentially also differentiate between those using ADM from those who do not.

Author contributions

Andreas Høstmælingen was responsible for designing and initiating the study, planning and performing the analyses, interpreting the results and writing the report, and was involved in managing the study and collecting the data.

Pål Gunnar Ulvenes was responsible for designing, initiating and managing the study, collecting the data, planning and performing the analyses, interpreting the results and was involved in writing the report.

Helene Amundsen Nissen-Lie was responsible for initiating and managing the study and was involved in designing the study, planning the analyses, performing the analyses, interpreting the results and writing the report.

Mikkel Eielsen was responsible for collecting the data and was involved in designing the study, planning and performing the analyses, interpreting the results and writing the report.

Bruce E. Wampold was involved in designing and initiating the study, interpreting the results and writing the report.

All authors approved the final version of the manuscript.

Role of the funding source

The study was funded by the University of Oslo and Modum Bad. The funder had no role in study design, data collection, -analysis, -interpretation, or writing of the report. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Statement of ethics

The study was conducted in compliance with APA ethical standards and IRB standards, and was reviewed and approved by the Norwegian regional committee for medical and health research ethics (application number 2014/2355 and 2016/2003).

Declaration of Competing Interest

All authors have completed the International Committee of Medical Journal Editors (ICMJE) form for disclosure of potential conflict of interest and declare that there is no conflict of interest.

Acknowledgements

None.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.jad.2021.01.066.

References

American Psychiatric Association, 2013. Diagnostic and statistical manual of mental disorders (DSM-5). American Psychiatric Publishing, Washington, DC.

American Psychiatric Association. (2010). Practice guideline for the treatment of patients with major depressive disorder. Retrieved from https://psychiatryonline.org/pb/assets/raw/sitewide/practice_guidelines/guidelines/mdd.pdf.

Asparouhov, T., Muthén, B.O., 2009. Exploratory Structural Equation Modeling. Structural Equation Modeling: A Multidisciplinary Journal 16 (3), 397–438.

Barnhofer, T., Crane, C., Hargus, E., Amarasinghe, M., Winder, R., Williams, J.M.G., 2009. Mindfulness-based cognitive therapy as a treatment for chronic depression: A preliminary study. Behaviour research and therapy 47, 366–373.

Beck, A.T., Steer, R.A., Brown, G.K., 1996. Beck depression inventory-II. Psychological Corporation, San Antonio, TX.

Belmaker, R.H., Agam, G., 2008. Mechanisms of disease: Major Depressive Disorder. The New England Journal of Medicine 358, 55–68.

Blatt, S.J., Quinlan, D.M., Chevron, E.S., McDonald, C., Zuroff, D., 1982. Dependency and self-criticism: Psychological dimensions of depression. Journal of Consulting and Clinical Psychology 50 (1), 113–124.

Bockting, C.L., Schene, A.H., Spinhoven, P., Koeter, M.W.J., Wouters, L.F., Huyser, J., Kamphuis, J.H., The DELTA Study Group, 2005. Preventing Relapse/Recurrence in Recurrent Depression With Cognitive Therapy: A Randomized Controlled Trial. Journal of Consulting and Clinical Psychology 73 (4), 647–657. https://doi.org/10.1037/0022-006X.73.4.647.

Borsboom, D., Cramer, A.O.J., 2013. Network Analysis: An Integrative Approach to the Structure of Psychopathology. Annual Review of Clinical Psychology 9, 91–121. https://doi.org/10.1146/annurev-clinpsy-050212-185608.

Boschloo, L., Bekhuis, E., Weitz, E.S., Reijnders, M., DeRubeis, R.J., Dimidjian, S., Dunner, D.L., Dunlop, B.W., Hegerl, U., Hollon, S.D., Jarret, R.B., Kennedy, S.H., Miranda, J., Mohr, D.C., Simons, A.D., Parker, G., Petrak, F., Herpertz, S., Quilty, L. C., Rush, A.J., Segal, Z.V., Vittengl, J.R., Schoevers, R.A., Cuijpers, P., 2019. The symptom-specific efficacy of antidepressant medication vs. cognitive behavioral therapy in the treatment of depression: results from an individual patient data meta-analysis. World Psychiatry 18 (2), 183–191.

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.

Brown, T.A., 2013. Latent Variable Measurement Models. In: Little, T.D. (Ed.), The Oxford Handbook of Quantitative Methods in Psychology: Vol. 2, Ed. Oxford University Press. Statistical Analysis.

Chekroud, A.M., Gueorguieva, R., Krumholz, H.M., Trivedi, M.H., Krystal, J.H., McCarthy, G., 2017. Reevaluating the Efficacy and Predictability of Antidepressant Treatments. A Symptom Clustering Approach. JAMA Psychiatry 74 (4), 370–378. https://doi.org/10.1001/jamapsychiatry.2017.0025.

Chen, F.F., 2007. Sensitivity of Goodness of Fit Indexes to Lack of Measurement Invariance. Structural Equation Modeling: A Multidisciplinary Journal 14 (3), 464–504.

Cheung, G.W., Rensvold, R.B., 2002. Evaluating Goodness-of-Fit Indexes for Testing Measurement Invariance. Structural Equation Modeling 9 (2), 233–255.

Craighead, W.E., Dunlop, B.W., 2014. Combination Psychotherapy and Antidepressant Medication Treatment for Depression: For Whom, When, and How. Annual review of psychology 65, 267–300.

de Jonge, M., Bockting, C.L., van Oppen, P., Van, H.L., Peen, J., Kikkert, M.J., Dekker, J. J.M., 2018. The association between the number of previous episodes and modifiable vulnerability factors in remitted patients with recurrent depression. PLoS ONE (11), 13. https://doi.org/10.1371/journal.pone.0206495.

Dent, J., Teasdale, J.D., 1988. Negative cognition and the persistence of depression. Journal of Abnormal Psychology 97 (1), 29–34.

DeRubeis, R.J., Cohen, Z.D., Forand, N.R., Fournier, J.C., Gelfand, L.A., Lorenzo-Luaces, L., 2014. The Personalized Advantage Index: Translating Research on Prediction into Individualized Treatment Recommendations. A Demonstration. PLoS ONE 9 (1). https://doi.org/10.1371/journal.pone.0083875.

DeRubeis, R.J., Zajecka, J., Shelton, R.C., Amsterdam, J.D., Fawcett, J., Xu, C., Young, P. R., Gallop, R., Hollon, S.D., 2020. Prevention of Recurrence After Recovery From a Major Depressive Episode With Antidepressant Medication Alone or in Combination With Cognitive Behavioral Therapy: Phase 2 of a 2-Phase Randomized Clinical Trial. JAMA Psychiatry 77 (3), 237–245. https://doi.org/10.1001/jamapsychiatry.2019.3900.

First, M.B., Gibbon, M., Spitzer, R.L., Williams, J.B.W., Benjamin, L.S, 1997. Structured Clinical Interview for DSM-IV Axis II Personality Disorders. American Psychiatric Press, Washington, DC.

Fried, E.I., Nesse, R.M., 2015. Depression sum-scores don't add up: why analyzing specific depression symptoms is essential. BMC Medicine 13. https://doi.org/10.1186/s12916-015-0325-4.

Gadermann, A.M., Alonso, J.A., Vilgaut, G., Zaslavsky, A.M., Kessler, R.C., 2012. Comorbidity and Disease Burden in the National Comorbidity Survey Replication (NCS-R). Depression and Anxiety 29, 797–806.

Hair, J.F., Tatham, R.L., Anderson, R.E., Black, W., 1998. Multivariate data analysis, Fifth ed. Prentice-Hall, London.

Hawley, L.L., Zuroff, D.C., Brozina, K., Ho, M.-H.R., Dobson, K.S., 2014. Self-Critical Perfectionism and Stress Reactivity Following Cognitive Behavioral Therapy for Depression. International Journal of Cognitive Therapy 7, 287–303.

Hoertel, N., Blanco, C., Oquendo, M.A., Wall, M.M., Olfson, M., Falissard, B., Franco, S., Peyre, H., Lemogne, C., Limosin, F., 2017. A comprehensive model of predictors of persistence and recurrence in adults with major depression: Results from a national 3–year prospective study. Journal of Psychiatric Research 95, 19–27.

Hollon, S.D., 2016. The efficacy and acceptability of psychological interventions for depression: where we are now and where we are going. Epidemiology and Psychiatric Sciences 25, 295–300. https://doi.org/10.1017/S2045796015000748.

Hollon, S.D., DeRubeis, R.J., Fawcett, J., Amsterdam, J.D., Shelton, R.C., Zajecka, J., Young, P.R., Gallop, R., 2014. Effect of Cognitive Therapy With Antidepressant Medications vs Antidepressants Alone on the Rate of Recovery in Major Depressive Disorder: A Randomized Clinical Trial. JAMA Psychiatry 71 (10), 1157–1164. https://doi.org/10.1001/jamapsychiatry.2014.1054.

Huang, C., Chen, J.-H., 2015. Meta-analysis of the Factor Structures of the Beck Depression Inventory-II. Assessment 22 (4), 459–472.

Humer, E., Kocsis-Bogar, K., Berger, T., Schröder, J., Späth, C., Meyer, B., Moritz, S., Lutz, W., Probst, T., Klein, J.P., 2020. A comparison of the three year course between chronic depression and depression with multiple vs. few prior episodes. Psychiatry Research 291. https://doi.org/10.1016/j.psychres.2020.113235.

Jobst, A., Brakemeier, E.-L., Buchheim, A., Caspar, F., Cuijpers, P., Ebmeier, K.P., Falkai, P., van der Gaag, R.J., Gaebel, W., Herpertz, S., Kurimay, T., Sabaß, L., Schnell, K., Schramm, E., Torrent, C., Wasserman, D., Wiersma, J., Padberg, F., 2016. European Psychiatric Association Guidance on psychotherapy in chronic depression across Europe. European Psychiatry 33, 18–36.

Kendler, K.S., Zachar, P., Craver, C., 2011. What kinds of things are psychiatric disorders? Psychological medicine 41, 1143–1150.

Kline, R.B., 2005. Principles and practices of structural equation modeling. Guilford, New York, NY.

Köhler, S., Chrysanthou, S., Guhn, A., Sterzer, P., 2019. Differences between chronic and nonchronic depression: Systematic review and implications for treatment. Depression and Anxiety 36, 18–30.

Lemmens, L.H.J.M., Müller, V.N.L.S., Arntz, A., Huibers, M.J.H, 2016. Mechanisms of change in psychotherapy for depression: An empirical update and evaluation of research aimed at identifying psychological mediators. Clinical Psychological Review 50, 95–107.

Liu, Y., Millsap, R.E., West, S.G., Tein, J.-Y., Tanaka, R., Grimm, K.J., 2017. Testing Measurement Invariance in Longitudinal Data With Ordered-Categorical Measures. Psychological methods 22 (3), 486–506.

Luyten, P., Sabbe, B., Blatt, S.J., Meganck, S., Jansen, B., De Grave, C., Maes, F., Corveleyn, J., 2007. Dependency and Self-Criticism: Relationship with Major Depressive Disorder, Severity of Depression, and Clinical Presentation. Depression and Anxiety 24 (8), 586–596.

Ma, S.H., Teasdale, J.D., 2004. Mindfulness-Based Cognitive Therapy for Depression: Replication and Exploration of Differential Relapse Prevention Effects. Journal of Consulting and Clinical Psychology 72 (1), 31–40.

Manian, N., Schmidt, E., Bornstein, M.H., Martinez, P., 2013. Factor structure and clinical utility of BDI-II factor scores in postpartum women. J Affect Disord 149, 259–268.

Marsh, H.W., Guo, J., Dicke, T., Parker, P.D., Craven, R.G., 2020. Confirmatory Factor Analysis (CFA), Exploratory Structural Equation Modeling (ESEM), and Set-ESEM: Optimal Balance Between Goodness of Fit and Parsimony. Multivariate Behavioral Research 55 (1), 102–119. https://doi.org/10.1080/00273171.2019.1602503.

Marsh, H.W., Lüdtke, O., Muthén, B., Asparouhov, T., Morin, A.J.S., Trautwein, U., Nagengast, B, 2010. A New Look at the Big Five Structure Through Exploratory Structural Equation Modeling. Psychological Assessment 22 (3), 471–491.

Marsh, H.W., Morin, A.J.S., Parker, P.D., Kaur, G, 2014. Exploratory Structural Equation Modeling: An Integration of the Best Features of Exploratory and Confirmatory Factor Analysis. The Annual Review of Clinical Psychology 10, 85–110. https://doi.org/10.1146/annurev-clinpsy-032813-153700.

Marsh, H.W., Muthén, B.O., Asparouhov, T., Lüdtke, O., Robitzsch, A., Morin, A.J.S., Trautwein, U, 2009. Exploratory structural equation modeling, integrating CFA and EFA: Application to students' evaluations of university teaching. Structural Equation Modeling: A Multidisciplinary Journal 16, 439–476.

Meredith, W., 1993. Measurement Invariance, Factor Analysis and Factorial Invariance. Psychometrica 58 (4), 525–543.

Meredith, W., Teresi, J.A., 2006. An Essay on Measurement and Factorial Invariance. Medical Care 44 (11), 69–77.

Millsap, R.E., 2011. Statistical approaches to measurement invariance. Routledge, New York, NY.

Mongrain, M., Leather, F., 2006. Immature Dependence and Self-Criticism Predict the Recurrence of Major Depression. Journal of Clincal Psychology 62 (6), 705–713.

Morin, A.J.S., Arens, A.K., Marsh, H.W, 2016. A Bifactor Exploratory Structural Equation Modeling Framework for the Identification of Distinct Sources of Construct-Relevant Psychometric Multidimensionality. Structural Equation Modeling: A Multidisciplinary Journal 23 (1), 116–139. https://doi.org/10.1080/10705511.2014.961800.

Morin, A.J.S., Marsh, H.W., Nagengast, B, 2013. Exploratory Structural Equation Modeling. In: Hancock, G.R., Mueller, R.O. (Eds.), Structural equation modeling: A second course, Eds. Information Age Publishing, Inc, Charlotte, NC. pp. 395–436.

Moussavi, S., Chatterji, S., Verdes, E., Tandon, A., Patel, V., Ustun, B., 2007. Depression, chronic diseases, and decrements in health: results from the World Health Surveys. The Lancet 370 (9590), 851–858.

Muthén, L.K., Muthén, B.O., 1998–2017. Mplus User's Guide, Eighth ed. Muthén & Muthén, Los Angeles, CA.

Neff, K.D., Vonk, R., 2009. Self-Compassion Versus Global Self-Esteem: Two Different Ways of Relating to Oneself. Journal of Personality 77 (1), 23–50. https://doi.org/10.1111/j.1467-6494.2008.00537.x.

Pearson, M., Brewin, C.R., Rhodes, J., McCarron, G., 2008. Frequency and Nature of Rumination in Chronic Depression: A Preliminary Study. Cognitive Behaviour Therapy 37 (3), 160–168.

Quilty, L.C., Zhang, K.A., Bagby, R.M., 2010. The latent symptom structure of the Beck Depression Inventory–II in outpatients with major depression. Psychological Assessment 22 (3), 603–608.

Reise, S.P., 2012. The Rediscovery of Bifactor Measurement Models. Multivariate Behavioral Research 47 (5), 667–696.

Rodriguez, A., Reise, S.P., Haviland, M.G., 2016. Evaluating Bifactor Models: Calculating and Interpreting Statistical Indices. Psychological methods 21 (2), 137–150.

Schweizer, K., 2010. Some Guidelines Concerning the Modeling of Traits and Abilities in Test Construction. European Journal of Psychological Assessment 26 (1), 1–2.

Shafer, A.B., 2006. Meta-analysis of the Factor Structures of Four Depression Questionnaires: Beck, CES-D, Hamilton, and Zung. Journal of Clinical Psychology 62 (1), 123–146.

Sheehan, D.V., Lecrubier, Y., Sheehan, K.H., Amorim, P., Janvas, J., Weiller, E., Hergueta, T., Baker, R., Dunbar, G.C., 1998. The Mini-International Neuropsychiatric Interview (M.I.N.I.): The Development and Validation of a Structured Diagnostic Psychiatric Interview for DSM-IV and ICD-10. Journal of Clinical Psychiatry 59 (suppl 20), 22–33.

Spielberger, C.D., Ritterband, L.M., Reheiser, E.C., Brunner, T.M., 2003. The nature and measurement of depression. International Journal of Clinical and Health Psychology 3 (2), 209–234.

Stewart, J.G., Harkness, K.L., 2012. Symptom specificity in the acute treatment of Major Depressive Disorder: A re-analysis of the treatment of depression collaborative research program. J Affect Disord 137, 87–97.

Subica, A.M., Fowler, J.C., Elhai, J.D., Frueh, B.C., Sharp, C., Kelly, E.L., Allen, J.G., 2014. Factor Structure and Diagnostic Validity of the Beck Depression Incentory–II With Adult Clinical Inpatients: Comparison to a Gold-Standard Diagnostic Interview. Psychological Assessment 26 (4), 1106–1115.

ten Have, M., de Graaf, R., van Dorsselaer, S., Tuithof, M., Kleinjan, M., Penninx, B.W.J. H., 2018. Recurrence and chronicity of major depressive disorder and their risk indicators in a population cohort. Acta Psychiatrica Scandinavica 137 (6), 503–515.

Vanheule, S., Desmet, M., Groenvynck, H., Rosseel, Y., Fontaine, J., 2008. The Factor Structure of the Beck Depression Inventory-II. An evaluation. Assessment 15 (2), 177–187.

Wang, Y.-P., Gorenstein, C., 2013. Psychometric properties of the Beck Depression Inventory-II: a comprehensive review. Brazilian Journal of Psychiatry 35 (4), 416–431.

Ward, L.C., 2006. Comparison of Factor Structure Models for the Beck Depression Inventory–II. Psychological Assessment 18 (1), 81–88.

Werner, A.M., Tibubos, A.N., Rohrmann, S., Reiss, N., 2019. The clinical trait self-criticism and its relation to psychopathology: A systematic review – Update. J Affect Disord 246, 530–547.

Zeeck, A., Wietersheim, J.v., Weiss, H., Hermann, S., Endorf, K., Lau, I., Hartmann, A., 2020. Self-Criticism and Personality Functioning Predict Patterns of Symptom Change in Major Depressive Disorder. Frontiers in Psychiatry 11. https://doi.org/10.3389/fpsyt.2020.00147.