Open Access

OBM Neurobiology

LIDSEN Publishing Inc.

Original Research

Assessing the Factor Structure and Construct Validity of the Beck Depression Inventory (BDI-II) in a Korean Preschool Teacher Sample

Yang Eun Kim [1, †], Boram Lee [2, †, *]

1. Department of Social Welfare, Woosong University, 171 Dongdaejeon-ro, Dong-gu, Daejoen, 34606, South Korea; E-Mail: yekim@wsu.ac.kr
2. Department of Early Childhood Education, Woosong University, 171 Dongdaejeon-ro, Dong-gu, Daejeon, 34606, South Korea; E-Mail: blee@wsu.ac.kr

† These authors contributed equally to this work.

* **Correspondence:** Boram Lee; E-Mail: blee@wsu.ac.kr

Academic Editors: Ines Testoni, Adriano Zamperini and Lorenza Palazzo

Special Issue: How COVID-19 Changed Individual and Social Life: Psychological and Mental illness Studies on the Pandemic Outcomes

OBM Neurobiology
2024, volume 8, issue 2
doi:10.21926/obm.neurobiol.2402224

Received: March 07, 2024
Accepted: May 14, 2024
Published: May 21, 2024

Abstract

Preschool teachers are at high risk of depression. Therefore, emphasis should be placed on evaluating a valid and reliable instrument to measure depressive symptomatology for this population. One such promising instrument is the Beck Depression Inventory (BDI-II), a 21-item self-report measure designed to assess the presence and severity of depression. Although the tool's psychometric properties have been analyzed in different samples, they have not been validated with a Korean preschool teacher sample. Hence, this study sought to evaluate the psychometric properties of the Korean version of the BDI-II in a sample of Korean preschool teachers. The study participants comprised 252 Korean preschool teachers (243 women and 9 men) with an average age of 33.5 years (SD = 10.3, range = 21-59). The participants completed the Korean version of the BDI-II online. Confirmatory factor analysis was performed to examine the psychometric properties of the BDI-II. Six models based on the

literature were tested, including single-factor, correlated two- and three-factor, and bifactor models. The reliability and validity of the BDI-II were also analyzed. The bifactor model demonstrated the best data fit with an overarching vital general depression factor that coexists with comparatively weak specific cognitive, affective, and somatic factors. Internal consistency, measured using McDonald's omega coefficients, was adequate in all cases. Using a validated screening instrument such as the BDI-II may allow clinicians to better detect depression among preschool teachers because of its conciseness and satisfactory psychometric properties. Since a bifactor model confirmed the unidimensionality of the BDI-II, the scores of specific factors should not be used in isolation. However, clinical assessment may benefit from its subscales if they are used in conjunction with total scores.

Keywords
Beck Depression Inventory-II; confirmatory factor analysis; convergent validity; depression; preschool teachers

1. Introduction

Depression is a common mood disorder that affects an individual's ability to perform life activities across different domains. Depression is so common and pervasive that it is often referred to as the "common cold" of psychological disorders [1]. More than 280 million people of all ages worldwide suffer from depression, significantly contributing to the global burden of disease [2]. In South Korea, depression is regarded as a significant public health concern as it is one of the leading causes of suicide death in the general population [3], among which teachers, particularly preschool teachers, are at a high risk. A cross-national comparison survey by the OECD [4] demonstrated that, compared to other countries, Korean teachers identify different sources of stress more often due to a high perceived volume of administrative work, time restraints, child behavior challenges, and addressing parents' concerns [5]. Moreover, preschool teachers specifically experience additional stressors when they respond to different emergencies involving young children and perform nonteaching-related tasks such as cleaning, dressing, and feeding—a result of working directly and intimately with young children [6]. Public awareness is also inadequate regarding the professionalism of Korean preschool teachers as essential partners of the preschool education profession, resulting in low social recognition of their status as well as measly salaries [7]. Furthermore, because the South Korean ECE system (e.g., curriculum, tuition, and operation) is centralized to the Ministry of Health and Welfare, teachers may experience increased stress levels as a result of administrative duties to meet government requirements [5]. Thus, their experience of such stressors increases their susceptibility to mental disorders, which in turn affects instructional quality and children's social and emotional development [8]. The growing prevalence of depressive symptoms among these professional educators has necessitated an accurate assessment of depression acuity and severity. Early detection allows for timely intervention and treatment, significantly improve an individual's well-being and long-term outcomes.

Among self-assessment instruments, the Beck Depression Inventory-II (BDI-II) is one of the most widely used to diagnose depression and assess its severity in adolescents and adults [9]. Since the

BDI-II's publication, its psychometric performance's robustness has been well established across diverse communities and populations [10-19]. Although the BDI-II was initially designed to map onto the Diagnostic and Statistical Manual of Mental Disorders symptom criteria for major depressive disorder, research suggests that the former is instead a measure of depression symptom severity, generally categorized as cognitive, affective, and somatic [20]. However, due to the original intent of the BDI-II to measure depression globally, existing studies on its factor structure have been highly controversial and inconclusive. In their initial study, Beck et al. concluded that BDI-II items assess the somatic-affective and cognitive dimensions of depression, which has been replicated by many studies [11, 12] but with specific differences regarding the extent to which the 21 items were associated with the two factors. Since then, several studies have shown different dimensions of the BDI-II structure using confirmatory factor analysis (CFA), mainly with two [13, 14] and three correlated factors [10, 15, 16]. For example, studies involving college student samples have supported an alternative two-factor model, but the factors were termed cognitive-affective and somatic [13, 14]. Meanwhile, other studies in primary care settings with cancer and heart disease patients, as well as substance abusers, observed that the three-factor model consisting of cognitive, somatic, and affective domains best represented the BDI-II's factor structure [10, 15, 16]. Furthermore, studies involving adolescents and undergraduate students have identified an alternative three-factor model of negative attitudes, performance difficulty, and somatic domains [17-19].

Although the above evidence supports two- or three-factor structures, researchers have suggested that first-order-factor solutions may be problematic as they fail to represent multidimensionality, which occurs when indicators are associated with more than one construct. Theoretically, the assumed multidimensionality of the BDI-II is not aligned with its current scoring system. A two- or three-factor structure of the BDI-II would theoretically translate into a separate calculation and interpretation of scores for each subscale. However, researchers typically calculate BDI-II total scores by aggregating across the 21 items, which suggests the presence of a general factor and contradicts the findings of factor analytic studies that BDI-II evaluates multiple depression domains [10-19]. However, studies have not clarified the degree to which the total scores reflect reliable variation on a single overall latent entity and yet concurrently reflect multiple subfactors of the same latent entity, the extent to which the use of subscale scores is justified [21].

Recent studies have addressed some of these limitations by testing and proposing alternative bifactor models [21-25], which partition item variance into an overarching latent construct and one of several specific content domains. Bifactor models can also be used to determine whether the scale items are primarily viewed as unidimensional and, therefore, whether the total score should be used or multidimensional, which requires a separate interpretation of subscales from the total score [23]. These studies showed that a bifactor model has a better data fit than alternative models. For example, Faro and Pereira [22] and Keller et al. [24] examined the BDI-II using bifactor modeling. They found that its subscale scores were unreliable, confirming the extremely low variance beyond that explained by the general factor. These findings are consistent with McElroy et al. [21], who also obtained similar results when they tested 15 different CFA models of the BDI-II derived from theory and other studies.

Although the psychometric properties of the BDI-II have been well documented among Korean populations, including adolescents [17], university students [26], and patients with major depression [27], whether the findings can be generalized to specific occupational groups such as

preschool teachers remain unclear. This is because the stress or depression experienced by preschool teachers is different from that observed among other populations (e.g., adolescents, older adults, and psychiatric patients). Hence, without a more rigorous empirical examination, one cannot assume that the measure would perform equally well in a teaching occupational group where the rates of severe depression have been found to be as high as 18.6% [28]. Because of the lack of consensus on the structure and dimensionality of the BDI and the symptom severity presented in this population, an examination of the measure's performance in a preschool sample would help understand the function of the BDI-II.

To address this literature gap, this study aimed to examine the factor structure of the BDI-II among preschool teachers. This will involve testing different measurement models suggested by other studies to determine which structural model best fits the data using CFA. Exploring the factor structure of the BDI-II among occupational groups in Korea may provide more extensive empirical data to improve the BDI-II. Next, this study also sought to examine the reliability and validity of BDI-II scores to establish the tool's psychometric robustness.

2. Materials and Methods

2.1 Participants and Procedures

This study obtained ethical approval from the Institutional Review Board at Woosong University in Korea, where the study took place (Protocol Code: 1041549-200457-SB-91). Afterward, a pilot study was conducted in which four teachers from childcare centers were recruited through the personal contacts of the principal investigator (PI), who visited the childcare centers in person to explain the purpose of the study and perform an initial assessment of the scale. Pilot participants were encouraged to suggest any necessary changes. No further changes or modifications were made, as the pilot participants indicated no changes were required. After the pilot study, an online link to the questionnaire was generated using Google Forms. Study participants were recruited through personal contacts, social media advertising, and postings on the Korean National Early Childhood Teachers' community website. These posts described the study's purpose, benefits, participation risks, voluntary nature, and data confidentiality.

Preschool teachers who met the following inclusion criteria were invited to participate: (a) born in the Republic of Korea, (b) employed by a childcare center (c) working as full- or part-time preschool teachers, and (d) willing to participate. They were not compensated for their participation. All interested participants were directed to the survey link to complete the questionnaire online. Informed consent was obtained from the participants electronically before their data were collected.

The full sample consisted of 252 preschool teachers (243 women and 9 men). Their mean age was 33.5 years (SD = 10.3, range = 21-59). The participants had an average of 6.3 years of teaching experience (SD = 6.02, range = 1 month-35 years). They worked at on-site childcare centers (49.6%), private childcare centers (21.0%), public childcare centers (11.9%), corporate childcare centers (6.7%), domestic daycare centers (5.6%), and others (5.2%).

Notably, gender imbalance in the early childhood workforce has always been a global phenomenon, where extensive studies have found that male early childhood educators lingered around 2%-3% in the majority of Western and non-Western countries [29]. No differences were observed in Korea, where only 3.4% of early childhood teachers are males. The latest data from the

Ministry of Health and Welfare show that an overwhelming 96.6% of preschool teachers are women [30], which explains why most of our sample were women.

2.2 Measure

2.2.1 The Beck Depression Inventory–Second Edition (BDI-II)

The BDI-II consists of 21 items on affective-cognitive and affective symptoms associated with depressive symptomatology. Each item is rated on a 4-point Likert-type scale ranging from 0 (symptom absent) to 3 (severe symptoms). Total severity scores range from 0 to 63, with higher scores indicating higher depression severity. Cutoff scores serve as guidelines for interpreting specific severity levels of depressive symptoms: minimal (0-13), mild (14-19), moderate (20-28), and severe (29-63). This study used the Korean version of the BDI-II, which has been validated in Korean adults [31].

2.3 Statistical Analysis

Statistical analysis was then performed using IBM SPSS Statistics for Windows (version 23) and Mplus 7.4. Before the analysis, the data for the 21 items of the BDI-II were screened for outliers, missing values, and normality. No outliers were removed because their percentage was considered small (0.6%) relative to the sample size. Missing data from participants who intentionally or unintentionally skipped or refused to answer some questions were minimal; the most significant number of missing cases was 11, less than 5% of the total cases in the dataset. These data were replaced using the expectation-maximization algorithm. The normality of the distribution was assessed through univariate analyses. As in many nonclinical samples, the data were non-normally distributed, and skewness and kurtosis values for all study variables were well within acceptable ranges (i.e., ±3.00).

Next, CFA via the weighted least squares estimator (WLSMV) was performed to assess the model that best fits the BDI-II. Studies have found that WLSMV estimation outperforms maximum likelihood for ordered categorical indicators involving five or fewer answer categories such as the BDI-II [32, 33]. Quality of fit was evaluated using chi-square (χ^2) and its subsequent ratio with degrees of freedom (χ^2/df), comparative fit index (CFI), Tucker–Lewis index (TLI), root mean square error of approximation (RMSEA), and its 90% confidence interval (90% CI). Marsh et al.'s guidelines suggest that the χ^2/df ratio should be less than 5; CFI and TLI ≥ 0.90, and RMSEA ≤ 0.06 to indicate a good fit [34].

Multiple CFAs were performed to identify six different latent structures of the BDI-II developed with adult samples and commonly used in the literature to test goodness of fit when applied to the data. Specifically, six models with the best fit from each validation study were selected. Model 1 was a one-factor solution in which all 21 items of the BDI-II were allowed to load onto a single latent depression variable. Model 2 tested Beck et al.'s original two-factor model, "somatic-affective" and "cognitive" factors. Model 3 was a correlated three-factor model postulated by Buckley et al. [10] consisting of cognitive, affective, and somatic factors. Model 4 was also a correlated three-factor model proposed by Wu and Huang [19], composed of negative attitude, performance difficulty, and somatic elements. This model also included three correlated item pairs (2 and 3; 4 and 12; 16 and 18). Model 5 was Faro and Pereira's [22] bifactor model consisting of a general factor and two

orthogonal group factors. All 21 items were loaded onto the general factor and two specific factors (cognitive-affective and somatic-affective). Finally, Model 6 was García-Barista et al.'s [23] bifactor model, which includes a general factor on which all items load and three specific factors (cognitive, affective, and somatic). (see Table 1).

Table 1 Factor models tested and factors on which items were mapped.

Model	Factors	Items
Model 1 (single factor)	General depression	1-21
Model 2 (Two-factor)	Somatic-affective,	4, 10-13, 15-21
	Cognitive	1-3, 5-9, 14
Model 3 (Three-factor)	Cognitive,	1-3, 5-9, 14
	Affective,	4, 10, 12, 13
	Somatic	11, 15-21
Model 4 (Three-factor)	Negative attitude,	1-3, 5-10, 14
	Performance difficulty,	4, 11-13, 17, 19
	Somatic elements	15, 16, 18, 20, 21
Model 5 (Bifactor)	General depression,	1-21
	Cognitive-affective,	1-10, 12, 14
	Somatic-affective	11, 13, 17, 19
Model 6 (Bifactor)	General depression,	1-21
	Cognitive,	3, 5-8, 13, 14
	Affective,	1, 2, 4, 9, 12
	Somatic	10, 11, 15-21

Reliability in terms of internal consistency was examined using McDonald's omega reliability coefficients, which have been proven to be more robust than Cronbach's alpha in capturing the proportion of scale variance due to all common factors and the proportion of scale variance attributable to a general factor [35] as is the case for the bifactor models tested in this study. Omega reliability coefficients (ω) estimate the degree of variance in total scale scores attributable to familiar sources of variables. In contrast, the omega hierarchy (ω*H*) refers to the estimated proportion of variance of the observed total scores that can be attributed solely to the general factor [36]. McDonald's omega ω values above 0.70 can be considered acceptable internal consistency estimates. In addition, ω*H* values above 0.80 suggest that factor strength and total scores can be considered unidimensional [37]. Additionally, explained common variance (ECV) was computed to interpret each factor's importance further. ECV assesses the proportion of the common variance explained by the general factor divided by the total common variance, indicating unidimensionality. An ECV value of >0.70 indicates a general solid factor and that the common variance is essentially unidimensional [38].

We performed an a priori power analysis using G Power 3.1 to determine the study's required sample size. For this sample, it was not appropriate to calculate margins of error because of the

nonrepresentative, nonrandom processes used to assemble it. Nevertheless, the current sample size 252 provides a power of 0.995, sufficient to detect a medium effect size of $f^2 = 0.15$.

3. Results

3.1 Descriptive Statistics

Table 2 shows that the 21 items of the BDI-II had an overall mean score of 14.8 (SD = 8.2). Based on the cutoff score guidelines, the participants scored in the mildly depressed range.

Table 2 Descriptive statistics for the Beck Depression Inventory-II items.

Item	*M*	*SD*	Range	Skewness	Kurtosis
1. Sadness	0.31	0.52	0-3	1.61	2.75
2. Pessimism	0.28	0.65	0-3	1.12	2.92
3. Past failure	0.40	0.79	0-3	1.65	1.19
4. Loss of pleasure	0.51	0.69	0-3	1.36	1.88
5. Guilty feelings	0.52	0.80	0-3	1.19	-0.44
6. Punishment feelings	0.28	0.65	0-3	1.87	1.81
7. Self-dislike	0.48	0.76	0-3	1.29	0.36
8. Self-criticalness	0.44	0.85	0-3	1.86	2.36
9. Suicidal thoughts	0.31	0.54	0-3	0.98	1.10
10. Crying	0.63	0.93	0-3	1.18	0.80
11. Agitation	0.46	0.89	0-3	1.24	0.67
12. Loss of interest	0.79	0.82	0-3	0.66	-0.46
13. Indecisiveness	1.11	1.21	0-3	0.60	-1.23
14. Worthlessness	0.30	0.66	0-3	1.21	1.73
15. Loss of energy	0.98	0.79	0-3	0.28	-0.76
16. Changes in sleeping	1.73	1.14	0-3	-0.35	-1.29
17. Irritability	0.62	0.77	0-3	1.13.	0.85
18. Changes in appetite	1.36	1.30	0-3	0.15	-1.73
19. Concentration difficulty	0.73	0.90	0-3	0.74	-0.94
20. Tiredness or fatigue	0.81	0.64	0-3	0.75	1.75
21. Loss of interest in sex	1.72	0.83	0-3	0.57	1.32
Overall mean score	14.8	8.2	0-63		

3.2 Model Comparisons

Table 3 summarizes the six tested CFA models and their respective fit indices. The one-factor model provided the poorest data fit, followed by the two- and three-factor correlated models, none of which reached an overall acceptable fit. In contrast, Faro and Pereira's bifactor model (model 5) fit adequately. In contrast, the bifactor model (model 6) consisting of a general depression factor and three specific factors (cognitive, affective, and somatic) demonstrated an overall good fit to the data with respect to CFI, TLI, and RMSEA statistics which were all in satisfactory to-good ranges fit (χ^2 = 239.1, df = 165, χ^2/df = 1.5, CFI = 0.93, TLI = 0.92, RMSEA = 0.042 (90% CI = 0.030-0.054)).

Therefore, this model had higher fit index values than all other tested models and was deemed the best fit. The bifactor model's suitability can be further determined by its parameter estimates.

Table 3 Confirmatory factor analysis: Fit indices for the alternative models of the BDI-II.

Model	Factors	χ^2	df	χ^2/df	CFI	TLI	RMSEA (90% CI)
Model 1	Single	383.3	189	2.0	0.80	0.80	0.064 (0.055-0.073)
Model 2	Two-factor	346.6	188	1.8	0.84	0.83	0.058 (0.048-0.067)
Model 3	Three-factor	314.9	186	1.7	0.87	0.85	0.053 (0.042-0.062)
Model 4	Three-factor	296.6	186	1.6	0.89	0.89	0.049 (0.038-0.059)
Model 5	Bifactor	256.4	167	1.5	0.91	0.90	0.046 (0.035-0.057)
Model 6	Bifactor	239.1	165	1.5	0.93	0.92	0.042 (0.030-0.054)

Notes: df = degrees of freedom; CFI = comparative fit index; TLI = Tucker-Lewis index; RMSEA = root mean square error of approximation; CI = confidence interval.

*$p < 0.01$.

The loadings on the general factor were all moderate to strong, ranging from 0.41 (item 13: indecisiveness) to 0.91 (item 3: past failure). Significantly, all items loaded higher on the general depression factor than on their respective specific factors, suggesting that the general factor explained a large proportion of the variance of the items. Meanwhile, the loadings on the specific factors (i.e., cognitive, affective, and somatic) were mostly weak; specifically, one somatic item (item 17: irritability) had extremely low loading on its specific factor, although significant. These results suggest that most variances of the cognitive, affective, and somatic items were almost completely explained by a general factor. Table 4 presents the standardized factor loadings of the bifactor model.

Table 4 Factor loadings of the BDI-II items for the bifactor model.

Item	General	Cognitive	Affective	Somatic
1. Sadness	0.75		0.39	
2. Pessimism	0.57		0.16	
3. Past failure	0.91	0.50		
4. Loss of pleasure	0.56		0.27	
5. Guilty feelings	0.77	0.20		
6. Punishment feelings	0.86	0.47		
7. Self-dislike	0.50	0.34		
8. Self-criticalness	0.82	0.33		
9. Suicidal thoughts	0.76		0.14	
10. Crying	0.80			0.31
11. Agitation	0.85			0.10
12. Loss of interest	0.74		0.11	
13. Indecisiveness	0.41	0.32		
14. Worthlessness	0.67	0.27		
15. Loss of energy	0.60			0.11
16. Changes in sleeping	0.86			0.47

17. Irritability	0.78	0.09
18. Changes in appetite	0.64	0.34
19. Concentration difficulty	0.89	0.16
20. Tiredness or fatigue	0.43	0.25
21. Loss of interest in sex	0.65	0.33

Note: All loadings $p < 0.01$

3.3 Bifactor Indices

The general depression factor had an ECV of 0.81, which explained the extracted common variance, while the specific cognitive, somatic, and affective factors accounted for the remaining 19%. Also, because this ECV value surpassed the 0.70 threshold, the model can be considered unidimensional, and the general factor of the BDI-II is the most appropriate indicator of depression. The omega reliability coefficient (ω) was 0.92, suggesting that the general and the three specific factors explained 92% of the total score variance. This indicates the presence of a reliable general factor that influences the variation across all items. Moreover, the omega hierarchical (ω*H*) was 0.88, showing that the general factor solely accounted for 88% of the variance in total scores and met the >70 criteria. Combined, the results of the ECV index and MacDonald's ω_H coefficients suggest that the general factor explained most of the variance.

Moreover, the ratio of the two indices (ω*H*/ω = 0.96) indicated that about half of the reliable variance in total scores can be attributed to the general factor. In contrast, only 4% (ω - ω*H* = 0.04) of the variance in total scores can be attributed to the multidimensionality due to the specific characteristics. Hence, raw total scores in the Korean version of the BDI-II can be interpreted as a unidimensional reflection of depression severity. Meanwhile, the omega hierarchical subscales were 0.24, 0.11, and 0.17 for the cognitive, affective, and somatic factors, respectively, indicating a low degree to which the BDI-II subscales reliably measure the specific variances of the three components.

4. Discussion

This study aimed to analyze the psychometric properties of the Korean version of the BDI-II in a sample of preschool teachers. Given the increased use of assessment tools for depression, it is vital to psychometrically evaluate instruments in terms of measure interpretation and scoring accuracy [21]. Considering other studies' empirical and theoretical arguments, we performed CFA on six competing models of the latent structure of the BDI-II, including one-factor, two-factor, three-factor, and bifactor models. We found confirmatory evidence for a bifactor model with a general depression factor and three specific factors (cognitive, affective, and somatic), suggesting that a single underlying dimension of depression drives the BDI-II. These results are consistent with those from studies that support the BDI-II's bifactor structure [21-24]. Moreover, the bifactor statistical indices calculated for the bifactor solution of the sample suggest that the BDI-II is a unidimensional scale. Our findings are aligned with a growing number of studies suggesting that the BDI-II measures a predominant general depression factor and three specific factors corresponding to the cognitive, affective, and somatic subscales. García-Barista et al. [23] found that the general factor in the bifactor model accounted for 78% of the common variance, while the three factors explained 22%.

Similar results were reported by McElroy et al. (ECV = 0.69) [21], Keller et al. (ECV = 0.86) [24], and Lim et al. (ECV = 0.81) [27].

The coefficient omega hierarchical (ω*H*) of the domain-specific factors in our study ranged from 0.11 to 0.24, while that of the general factor remained high at 0.88. The corresponding values for the adult samples in McElroy et al. were 0.84 for the general factor alone and 0.06, 0.01, and 0.01 for the three specific factors, respectively. All these indices strongly support the unidimensionality of the BDI-II.

However, studies supporting the bifactor model have shown different model structures with two specific factors: cognitive-affective and somatic. Wang et al. found that the BDI-II was better explained by a bifactor model (cognitive-affective and somatic) in a middle school teacher sample from Mainland China [39]. The same factor structure was observed by Brouwer et al., Dere et al., and Faro and Pereira in a sample of clinical outpatients, Chinese- and European-heritage college students, and community-dwelling adults in Brazil, respectively [18, 40, 41]. These variations in findings may be partly explained by the recruitment of specific subgroups (e.g., Brazilian adults and Chinse adults), differences in age ranges (e.g., older adults vs. college students), participant characteristics (e.g., clinical vs. community-based participants), and sample size.

Analysis of the distributions of depression diagnoses showed that the average BDI-II total score for our participants was higher than those obtained by Yu et al. [26] for Korean university students (*M* = 10.63; *SD* = 8.8) and Lim et al. [27] for Korean nonclinical adults (*M* = 10.62; *SD* = 7.0) but significantly lower than the mean reported for a clinical sample with major depressive disorder (*M* = 30.4; *SD* = 11.59) [27].

Concerning the scoring of the BDI-II, all items loaded significantly onto a general depression factor, and their factor loadings were higher on the general factor compared with the domain-specific factors. Therefore, the strength of the general depression factor, as evidenced by ω and ω*H* and solid loadings of the general factor on the items, justifies the use and interpretation of a total depression score, with higher totals reflecting greater depression severity. Although this study highly emphasized the overall importance of the general depression factor, it did not completely invalidate the use of subscales in all contexts. Our results prove that although a substantial proportion of the variance in BDI-II scores is attributable to a general factor, the three specific factors (cognitive, affective, and somatic) did explain a nonredundant variance. Research has indicated the differential responses of various subfactors of depression symptoms to treatment [42]. Although the validity of the isolated use of subscales may be questioned [21, 43], this could be done in conjunction with total scores in treatment decisions [21, 22]. This is a significant outcome and supports the original instructions of Beck et al. [9] regarding how the measure should be scored.

With the high prevalence of depression among preschool teachers and its associated adverse outcomes, the BDI-II may be a valuable tool across interventions and prevention strategies. We believe that our findings, particularly those regarding the latent structure of the BDI-II, will allow for early diagnosis and treatment of preschool teachers at high risk of depression, thus helping healthcare professionals implement necessary and timely interventions against such disorders.

Notwithstanding the aforementioned implications, this study has several strengths and limitations. The strengths of this study include its use of a relatively understudied sample of preschool teachers in Korea and its performance of CFA on a range of competing models. Through CFA, this study extends prior work by replicating the BDI-II's factor structure to preschool teachers who are more vulnerable to depression. The study also adds to the growing evidence of BDI-II as a

valuable measure of depression severity for use with preschool teachers. The results also support a bifactor structure of the BDI-II and demonstrate adequate internal consistency.

Regarding weaknesses, most of the participants were female preschool teachers; hence, the sample does not represent the Korean population. Although women teachers are overrepresented mainly in early childhood education, more rigorous studies must include a more prominent representative and balanced sample to confirm the current factorial structure of the BDI-II. Second, teachers are a specific population that is vulnerable to and has a high risk of mental health conditions [5, 6, 44]. However, the participants in this study scored within the mildly depressed range, suggesting that while they have low mood and other symptoms of depression, such symptoms are less intense. Nevertheless, the BDI-II structure may differ between a sample of the general population and that of different occupational groups, and the findings cannot be generalized to them. Hence, it would be worthwhile to compare the structure of the BDI-II between nonclinical and clinical samples of preschool teachers to provide a clearer understanding of depression symptomatology among this population and determine whether the results can be replicated to demonstrate how the BDI-II can be generalized across different occupations. Next, this study is constrained by its smaller sample size compared with other validation studies performing CFA. Some studies that have conducted CFA with larger samples have increased CFI, TLI, and GFI cutoff values to 0.95 as evidence of a good model fit. Given this study's smaller sample size, the lower cutoff values (0.90) based on Marsh et al.'s recommendations have been chosen as evidence of acceptable model fit. Nevertheless, the models should be tested in a larger sample to further support the factor structure of the BDI-II in a preschool population. Another potential limitation is the cross-sectional data, which prevented us from evaluating the psychometric stability of the BDI-II structure over time. Hence, future studies may consider longitudinal designs across the lifespan. Data regarding the participants' psychiatric backgrounds were not obtained. This is an important variable as the literature suggests that the BDI-II model may vary across different samples (e.g., depressed vs. nondepressed) [22]. Therefore, future research would benefit from including participants' medical or clinical history to further examine the BDI-II's item structure and provide additional evidence of its capacity to assess depressive symptomatology. Finally, this study did not examine the relations between BDI-II scores and convergent measures. Therefore, the convergent validity of the Korean version of the BDI-II should be confirmed by examining its relations to other depression scales.

In conclusion, the results suggest that the Korean version of the BDI-II displayed acceptable factorial validity, best represented by a bifactor model consisting of a typical general depression factor and three specific factors (cognitive, affective, and somatic). Bifactor statistical indices also show that the BDI-II captures overall depressive severity, thus reflecting a unidimensional scale. Moreover, the findings further contribute to the potential utility of the scale's total score as a depression measure. While the scores of the three specific factors should not be used as the first choice, it would be more appropriate to use BDI-II total scores and scores corresponding to each subscale for both statistical and clinical purposes.

Author Contributions

Conceptualization, B.L.; methodology, B.L.; formal analysis, B.L.; investigation, Y.K.; data curation, Y.K.; writing—original draft preparation, B.L.; writing—review and editing, visualization, B.L.;

supervision, Y.K.; project administration, Y.K. All authors have read and agreed to the published version of the manuscript.

Funding

This research was supported by Woosong University.

Competing Interests

The authors have declared that no competing interests exist.

References

1. Dehestani M, Ghanbarian Sharabiani S, Nooripour R, Zanganeh F. Prediction of depression symptoms based on personality traits and romantic relationships among students. J Res Health. 2018; 8: 173-181.
2. World Health Organization. Depression disorder (depression). Geneva: World Health Organization; 2023.
3. Lee K, Lee O. Depression and suicidal ideation in patients with mental illness in South Korea: The mediating effect of alcohol drinking. Healthcare. 2023; 11: 2711.
4. OECD. Building a high-quality early childhood education and care Workforce: Further results from the Starting Strong Survey 2018 [Internet]. Paris: OECD; 2020. Available from: https://www.oecd-ilibrary.org/education/building-a-high-quality-early-childhood-education-and-care-workforce_b90bba3d-en.
5. Byun S, Jeon L. Early childhood teachers' work environment, perceived personal stress, and professional commitment in South Korea. Child Youth Care Forum. 2023; 52: 1019-1039.
6. Wiltshire CA. Early childhood education teacher workforce: Stress in relation to identity and choices. Early Child Educ J. 2024; 52: 655-668.
7. Ahn NR, Kim HS, Ahn SH. The relationship between job stress and turnover intention of child care teachers and the moderating role of motivation for child care work. J Fam Batter Life. 2015; 33: 87-102.
8. Ju C, Lan J, Li Y, Feng W, You X. The mediating role of workplace social support on the relationship between trait emotional intelligence and teacher burnout. Teach Teach Educ. 2015; 51: 58-67.
9. Beck AT, Steer RA, Brown GK. Manual for the Beck Depression Inventory-II. San Antonio, TX: Psychological Corporation; 1996.
10. Buckley TC, Parker JD, Heggie J. A psychometric evaluation of the BDI-II in treatment-seeking substance abusers. J Subst Abuse Treat. 2001; 20: 197-204.
11. Campos RC, Goncalves B. The Portuguese version of the Beck Depression Inventory-II (BDI-II): Preliminary psychometric data with two nonclinical samples. Eur J Psychol Assess. 2011; 27: 258-264.
12. Wang Z, Yuan CM, Huang J, Li ZZ, Chen J, Zhang HY, et al. Reliability and validity of the Chinese version of the beck depression inventory-II among patients with depression. Chin Ment Health J. 2011; 25: 476-480.

13. Whisman MA, Perez JE, Ramel W. Factor structure of the Beck Depression Inventory-Second Edition (BDI-II) in a student sample. J Clin Psychol. 2000; 56: 545-551.
14. Yang WH, Liu SL, Zhou T. Reliability and validity of the Chinese version of the Beck Depression Inventory-II in Chinese adolescents. Chin J Clin Psychol. 2014; 22: 240-245.
15. Almeida S, Camacho M, Barahona-Corrêa JB, Oliveira J, Lemos R, da Silva DR, et al. Criterion and construct validity of the Beck Depression Inventory (BDI-II) to measure depression in patients with cancer: The contribution of somatic items. Int J Clin Health Psychol. 2023; 23: 100350.
16. Ginting H, Näring G, van der Veld WM, Srisayekti W, Eni S, Becker ES. Validating the Beck Depression Inventory-II in Indonesia's general population and coronary heart disease patients. Int J Clin Health Psychol. 2013; 13: 235-242.
17. Lee EH, Lee SJ, Hwang ST, Hong SH, Kim JH. Reliability and validity of the Beck Depression Inventory-II among Korean adolescents. Psychiatry Investig. 2017; 14: 30-36.
18. Osman A, Kopper BA, Barrios F, Gutierrez PM, Bagge CL. Reliability and validity of the Beck depression inventory--II with adolescent psychiatric inpatients. Psychol Assess. 2004; 16: 120-132.
19. Wu PC, Huang TW. Gender-related invariance of the Beck Depression Inventory II for Taiwanese adolescent samples. Assessment. 2014; 21: 218-226.
20. Nejati S, Ariai N, Björkelund C, Skoglund I, Petersson EL, Augustsson P, et al. Correspondence between the neuropsychiatric interview M.I.N.I. and the BDI-II and MADRS-S self-rating instruments as diagnostic tools in primary care patients with depression. Int J Gen Med. 2020; 13: 177-183.
21. McElroy E, Casey P, Adamson G, Filippopoulos P, Shevlin M. A comprehensive analysis of the factor structure of the Beck Depression Inventory-II in a sample of outpatients with adjustment disorder and depressive episode. Ir J Psychol Med. 2018; 35: 53-61.
22. Faro A, Pereira CR. Factor structure and gender invariance of the Beck Depression Inventory - Second edition (BDI-II) in a community-dwelling sample of adults. Health Psychol Behav Med. 2020; 8: 16-31.
23. García-Batista ZE, Guerra-Peña K, Cano-Vindel A, Herrera-Martínez SX, Medrano LA. Validity and reliability of the Beck Depression Inventory (BDI-II) in general and hospital population of Dominican Republic. PLoS One. 2018; 13: e0199750.
24. Keller F, Kirschbaum-Lesch I, Straub J. Factor structure and measurement invariance across gender of the Beck Depression Inventory-II in adolescent psychiatric patients. Front Psychiatry. 2020; 11: 527559.
25. Maggi G, D'Iorio A, Aiello EN, Poletti B, Ticozzi N, Silani V, et al. Psychometrics and diagnostics of the Italian version of the Beck Depression Inventory-II (BDI-II) in Parkinson's disease. Neurol Sci. 2023; 44: 1607-1612.
26. Yu B, Lee HK, Lee K. Validation and factor structure of Korean version of the Beck Depression Inventory Second Edition (BDI-II): In a university student sample. Korean J Biol Psychiatry. 2011; 18: 126-133.
27. Lim SY, Lee EJ, Jeong SW, Kim HC, Jeong CH, Jeon TY, et al. The validation study of Beck Depression Scale 2 in Korean version. Anxiety Mood. 2011; 7: 48-53.
28. Ministry of Health and Welfare. Survey on the current status of nationwide childcare in 2021 [Internet]. Seoul: Korea Institute of Child Care and Education; 2021. Available from: https://repo.kicce.re.kr/handle/2019.oak/5300.

29. Reich-Shapiro M, Cole K, Plaisir JY. "I am the teacher": How male educators conceptualize their impact on the early childhood classroom. J Early Child Teach Educ. 2021; 42: 381-403.
30. Ministry of Health and Welfare. Childcare statistics 2023 [Internet]. Seoul: The Office for Korea Childcare Promotion Institute. Available from: https://www.kcpi.or.kr/kcpi/cyberpr/childcarestat1.do.
31. Sung HM, Kim JB, Park YN, Bai DS, Lee SH, Ahn HN. A study on the reliability and the validity of Korean version of the Beck Depression Inventory-II (BDI-II). J Korean Soc Biol Ther Psychiatry. 2008; 14: 201-212.
32. Bandalos DL. Relative performance of categorical diagonally weighted least squares and robust maximum likelihood estimation. Struct Equ Modeling. 2014; 21: 102-116.
33. Finney SJ, DiStefano C. Nonnormal and categorical data in structural equation models. In: A second course in structural equation modeling. Greenwich, CT: Information Age; 2006. pp. 269-314.
34. Marsh HW, Hau KT, Wen Z. In search of golden rules: Comment on hypothesis testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu & Bentler's (1999) findings. Struct Equ Modeling. 2004; 11: 320-341.
35. Zinbarg RE, Revelle W, Yovel I, Li W. Cronbach's α, Revelle's β, and McDonald's ω_H: Their relations with each other and two alternative conceptualizations of reliability. Psychometrika. 2005; 70: 123-133.
36. Rodriguez A, Reise SP, Haviland MG. Evaluating bifactor models: Calculating and interpreting statistical indices. Psychol Methods. 2016; 21: 137-150.
37. Cortina JM. What is coefficient alpha? An examination of theory and applications. J Appl Psychol. 1993; 78: 98-104.
38. Reise SP. The rediscovery of bifactor measurement models. Multivar Behav Res. 2012; 47: 667-696.
39. Wang X, Wang Y, Xin T. The psychometric properties of the Chinese version of the Beck Depression Inventory-II with middle school teachers. Front Psychol. 2020; 11: 548965.
40. Brouwer D, Meijer RR, Zevalkink J. On the factor structure of the Beck Depression Inventory-II: G is the key. Psychol Assess. 2013; 25: 136-145.
41. Dere J, Watters CA, Yu SC, Bagby RM, Ryder AG, Harkness KL. Cross-cultural examination of measurement invariance of the Beck Depression Inventory-II. Psychol Assess. 2015; 27: 68-81.
42. Mallinckrodt CH, Prakash A, Houston JP, Swindle R, Detke MJ, Fava M. Differential antidepressant symptom efficacy: Placebo-controlled comparisons of duloxetine and SSRIs (fluoxetine, paroxetine, escitalopram). Neuropsychobiology. 2007; 56: 73-85.
43. Subica AM, Fowler JC, Elhai JD, Frueh BC, Sharp C, Kelly EL, et al. Factor structure and diagnostic validity of the Beck Depression Inventory-II with adult clinical inpatients: Comparison to a gold-standard diagnostic interview. Psychol Assess. 2014; 26: 1106-1115.
44. Deguchi Y, Iwasaki S, Kanchika M, Nitta T, Mitake T, Nogi Y, et al. Gender differences in the relationships between perceived individual-level occupational stress and hazardous alcohol consumption among Japanese teachers: A cross-sectional study. PLoS One. 2018; 13: e0204248.