

Psychometric testing of the Indonesian version of beck depression inventory-II among Indonesian floods survivors

Tutu April Ariani^{1,3*}, Anastasia Anna², Henik Tri Rahayu³, Nur Aini³, Heni Dwi Windarwati⁴, Taty Hernawaty², Sriyani Padmalatha Konara Mudiyansele^{5,6}, and Mei-Feng Lin⁷

¹ Department of Nursing, College of Medicine, National Cheng Kung University, Tainan, Taiwan

² Faculty of Nursing, Padjadjaran University, Bandung, Indonesia

³ Department of Nursing, Faculty of Health Sciences, University of Muhammadiyah Malang, Malang, Indonesia

⁴ Department of Mental Health Nursing, Faculty of Health Sciences, Universitas Brawijaya, Malang, Indonesia

⁵ Operation Theatre Department, The National Hospital of Sri Lanka, Colombo, Sri Lanka

⁶ Nursing Faculty, University of Colombo, Colombo, Sri Lanka

⁷ Department of Nursing, College of Medicine, National Cheng Kung University, Tainan, Taiwan

*Correspondence: Mei-Feng Lin. Address: Department of Nursing, College of Medicine, National Cheng Kung University, Tainan, Taiwan. Email: l30mf@mail.ncku.edu.tw

Responsible Editor: Retnayu Pradanie

Received: 5 July 2023 ◦ Revised: 24 September 2023 ◦ Accepted: 26 September 2023

ABSTRACT

Introduction: Indonesia is a multi-linguistic country using the official language of Bahasa Indonesia. It is important to use formal language to collect research data without misinterpretation of study outcome and intervention effect with high accuracy. Therefore, this study aimed to analyze whether the Beck Depression Inventory-II (BDI-II) instrument translated into the Indonesian version was reliable and valid for measuring depression in floods-affected communities.

Methods: The forward-backward translation was used to translate the BDI-II from English to Indonesian version. We tested the reliability and validity including content and construct validity through exploratory factor analysis (EFA) with varimax rotation and confirmatory factor analysis (CFA). We recruited 107 annual flood survivors in West Java, Indonesia as participants to test the psychometric testing.

Results: The structures shown by the EFA are two factors, with Factor 1 representing negative cognitive-attitude domains and somatic symptom items consisting of Factor 2. The CFA concludes that the general factor model best fits the data on the basis Goodness-of-Fit Index (GFI)= 0.8; Root Mean Square Error of Approximation (RMSEA)= 0.09; Standard Root-Mean-Square Residual (SRMSR)); Comparative Fit Index (CFI)= 0.81; Tucker–Lewis Index (TLI)= 0.79; and p-value $\chi^2 = < 0.01$, χ^2/df : 1.82).

Conclusions: The results showed that the BDI-II Indonesian version has good reliability and validity and can be used to measure depression status among people who suffer floods in community setting. Future studies need to be validated within multiple socio-cultural groups.

Keywords: bahasa Indonesia, community, depression instrument, floods disaster, reliability and validity

Introduction

Indonesia was ranked 37th out of 180 countries most at risk of experiencing disasters in 2019 based on the World Risk Index. One of the major disasters in Indonesia is flooding, which recurs from year to year. Referring to data from the Indonesian National Disaster Management Agency (2023), throughout 2022, of the eight types of natural disasters in Indonesia, floods are among the most frequent. Floods are the most common of the eight types of events by 1520 times. Flood is a natural force that occurs and causes many positive and negative impacts; the negative impacts are often fatal. The adverse effect often pushed on mental health is depression compared to post traumatic stress disorder and anxiety (TellesSingh and Joshi, [2009](#); Mason, Andrews and Upton, [2010](#); Wind et al., [2013](#); Zhong et al., [2018](#)).

Previous studies have confirmed that when depression occurs, it will not be immediately obvious and is not treated correctly (Cacheda et al., [2019](#)). A review study by Penninx et al. ([2013](#)) concluded that depression could cause somatic disorders in a person, namely metabolic syndrome, inflammation up-regulations, and hypercortisolemia. In addition, a meta-analysis study showed that depression might increase the risk of coronary heart disease (CHD) (Gan et al., [2014](#)). In the social aspect, depression can be more sensitive to social rejection, acceptance, and negative social interaction problems (Steger and Kashdan, [2009](#)). Considering the high risk of depression, we need a valid and reliable measurement tool for early detection to measure the severity of depression in various settings, including community settings affected by flooding.

In the past few decades, the Beck Depression Inventory-II (BDI-II) has become one of the most commonly used measures to assess the symptoms and severity of depression in adolescents and adults. A comprehensive review by Wang and Gorenstein ([2013](#)) of 118 articles translated into 17 languages have been used in various countries in Europe, the Middle East, Asia, and Latin America in three setting samples, namely non-clinical (n = 47), psychiatric/ institutionalized (n = 37); and medical samples (n = 34). However, none of them mentioned the conversion of BDI-II into Indonesian. While many studies have established the psychometric properties of BDI-II worldwide, we have not found yet studies which have investigated its validity and reliability in Indonesia (Wang and Gorenstein, [2013](#)). Apart from this review, Ginting et al. ([2013](#)) conducted a study to test the validity and determined

the cut-off point for the BDI-II from the English to Bahasa Indonesia version among Indonesians. Unfortunately, the test was conducted on generally healthy participants, CHD, and depressed patients. Few published studies have been conducted in Indonesia using the BDI-II to measure depression in the community; however, they did not include flood victims in specific areas or mention a translation version of the Bahasa Indonesia with proper reliability and validity testing (Bei et al., [2013](#)).

Nevertheless, Indonesia is a multi-linguistic country using the official language of Bahasa Indonesia. Therefore, more people can understand the meaning of questionnaire content without misinterpretations. Also, research result needs to be accurate with the results of depression experienced by the participants and this will help to minimize an erroneous treatment of depression. Consequently, the Indonesian version of the BDI-II is needed after going through the critical measurement stages so that it is understandable and easy to fill in and can be used as a measuring tool for the right level of depression. In addition, for Indonesia's population, which ranks fourth in the world and is spread all over the world, it is significant to do a translation of the Indonesian version, which can be not only applied in the Indonesia community itself but also to immigrants around the world who can use it because of the floods. Indeed, the cultural background is still a limitation to be used as a generalization and is always an exciting topic in the discussion section (Ghassemzadeh et al., [2005](#); Wang and Gorenstein, [2013](#); González, Reséndiz and Reyes-Lagunes, [2015](#)).

Indonesia has a diverse culture that remains rooted in local wisdom in coping with natural disasters, including the psychological aftermath (Kadiyono and Harding, [2017](#); Agusintadewi, [2019](#); Samson et al., [2021](#)). Also, data show the largest annual floods occur in the Java Island (BNPB, 2023). Cultural diversity is more dominant to Sundanese, Javanese, etc. The Sundanese community has a custom of "Balai" which consists of three elements, namely Larangan (custom), Paharaman (religion) and Harim (state) to prepare themselves from natural disasters (Samson et al., 2021). However, there is a gap in the standardized depression measurement scale and traditional methods on accuracy of depression diagnosis and also instruments are user friendly for healthcare professional and the victims. Therefore, this study aimed to analyze whether the instrument translated into the Indonesian version was reliable and valid for measuring depression in flood-affected communities. There were two main goals of this study.

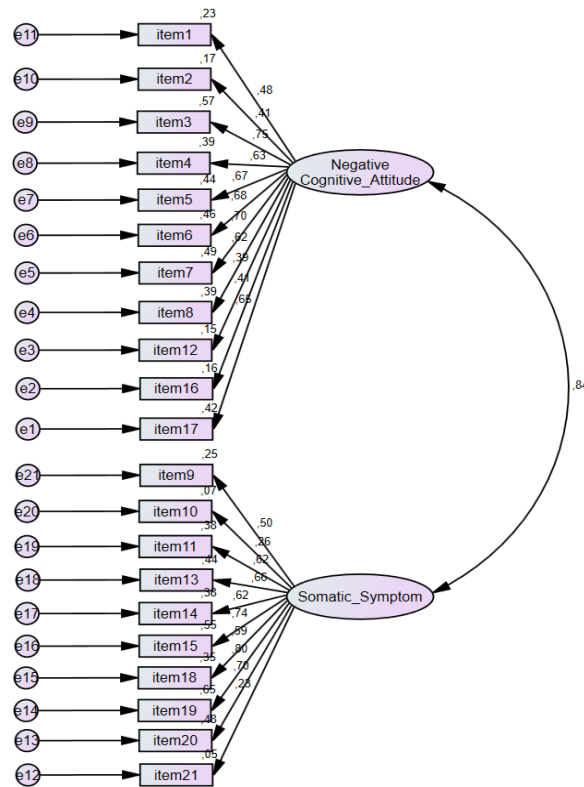


Figure 1. Schematic representation of two-factor model based on the CFA analysis of the BDI-II-Indonesian version
 Note: p-value $\chi^2 = < 0.01$, χ^2/df : 1.82, GFI= 0.8, RMSEA= 0.09, SRMR= 0.07, AGFI= 0.72, TLI= 0.79), IFI= 0.82, NFI= 0.67, CFI= 0.81, PNFI= 0.72, PGFI= 0.63

First, this study addressed to translate the BDI-II English version into Indonesian. The second aim was psychometric testing of the translation.

Materials and Methods

Instrument

The BDI– Second Edition is a 21-item measure of depression that was revised to include DSM-IV symptoms of depression—equivalent to DSM-5 symptoms—and different cognitive symptoms of depression (Ghassemzadeh et al., 2005; Ginting et al., 2013). Furthermore, the BDI-II is a 21-item self-report questionnaire for measuring the severity of depression in adolescents and adults, but not as a diagnostic tool. The BDI-II was revised in 1996 to be more consistent with DSM-IV criteria for depression. Individuals may rate their responses to items on a 0-3 scale, and the total scores range from 0-63 with the following cut-offs: 0-13: minimally depressed; 14-19: mildly depressed; 20-28: moderately depressed; and 29-63: severely depressed.

Forward and Backward BDI-II Translation

Before we carried out the BDI-II translation process, we asked permission to the developer by email to translate the instrument into an Indonesian version. The translation process consists of three stages including: 1) an independent professional English translator

translated the original BDI-II into Indonesian, 2) another professional translator from a language institution translated back the Indonesian version into English, 3) an expert in Mental Health Nursing compared the translated English version with the original and reviewed the translation to ensure the accuracy of the Indonesian language structure that was easy to understand by Indonesian participants. The research team continuously followed up and maintain progress of the discussion the during translation process to ensure consistency.

Psychometric testing of the BDI-II Indonesian version

Participant and recruitment procedure

After obtaining approval from the Health Research Ethics Committee, University of Muhammadiyah Malang (No.E.5. a/094/KEPK-UMM/V/2021), we recruited 107 people who speak Bahasa Indonesia and live in areas with continuous annual floods in West Java, Indonesia, using the 5:1 subject-to-item ratio sample size criterion (Gorsuch, 1988) to test the reliability of the BDI-II Indonesian version through a personal approach. Then, people who had the potential to become respondents provided their contact numbers as a tool to communicate during data collection. Participants were briefed about the purpose of the study and what they needed to do. Meanwhile, we explained the benefits

and possibilities of what is experienced after filling the inventory because this survey deals with psychological problems. Participants were explained as to what symptoms needed to be reported and how to apply the anxiety management protocol provided by the researcher. If the level of anxiety increased, then they could consult the closest mental health professional to get treatment. Finally, survey participants would receive a free voucher compensation on their cell phones. We excluded respondents who had lived less than three years at the research site, were not able to fill out the BDI-II through the Google application form or could not complete it.

Reliability testing

We tested internal consistency by using Cronbach's alpha to test the reliability of the Indonesian version of the BDI-II scale. Cronbach's alpha value greater than 0.90 indicates excellent reliability, and a value ≥ 0.70 indicates adequate internal consistency. Meanwhile, we tested three aspects of validity: content validity, convergent, and discriminant validity. Additionally, we also tested the inter-item correlation and item-total correlation. Kellar and Kelvin (2013) mentioned that 0.3-0.7 inter-item correlations were acceptable, while item-to-total correlations greater than 0.50 were considered satisfactory (Nawi et al., 2020).

Content validity testing

Researchers started testing content validity through the following steps: we asked eight Mental Health Nursing experts with more than five years of work experience, either as Mental Health Practitioners or lecturers in Indonesia, to assess the BDI-II Indonesian version for content validity, including relevance and clarity of the 21 questions. To assess the item relevancy, the eight experts were asked to rate each item using a four-point scale: not relevant (1), somewhat relevant (2), quite relevant (3), and highly relevant (4). Meanwhile, they used a three-point scale for clarity: very clear (3), item needs some revision (2), and not clear (1).

Then, we calculated the Item level of content validity index (I-CVI) for each item as the number of experts giving a rating of either 4 or 3 divided by the number of experts—the proportion in agreement about relevance and clarity. The values can range from 0 to 1, where an I-CVI > 0.79 indicates the item is relevant, an I-CVI between 0.70 and 0.79 indicates the item needs to be revised, and an I-CVI below 0.70 indicates the item should be eliminated (Rodrigues et al., 2017). Meanwhile, we computed the scale level of the content

Table 1. Demographic characteristics of participants

Characteristics	Participant n=107	
	M (SD)	n (%)
Age	32.19(\pm 13.86)	
Gender		
Male	35(32.71)	
Female	72(67.29)	
Marital status		
Married	83(77.57)	
Single	22(20.56)	
Divorced	2(1.87)	
Employment status		
Employee	16(14.95)	
Worker	15(14.01)	
Unemployed	76(71.03)	
Education		
Lower Secondary	39(36.45)	
Upper Secondary	61(57)	
Tertiary	7(6.54)	
Level of Depression		
Low		
Normal ups and down	64 (59.8)	
Mild mood disturbance	29 (27.1)	
Moderate		
Borderline clinical depression	1 (9)	
Moderate depression	7 (6.5)	
Significant		
Severe depression	5 (4.7)	
Extreme depression	1 (0.9)	

validity index (S-CVI) using two methods: the universal agreement (UA) among experts (S-CVI/UA) and the average content validity index (S-CVI/Ave). The S-CVI/UA was calculated by summing all items with I-CVI equal to 1 divided by the number of items, whereas the S-CVI/Ave was calculated by dividing the total I-CVI by the number of items. S-CVI/UA ≥ 0.8 and S-CVI/Ave ≥ 0.9 indicate that the items have excellent content validity (Zamanzadeh et al., 2015).

Construct validity testing

We tested construct validity using exploratory factor analysis (EFA) and confirmatory factor analysis (CFA). Since the normal distribution was violated and following the recommendation from El-Den et al. (2018), we used the principal factorial axis (PFA) extraction method and varimax as the rotation method in the EFA. Varimax was the most popular orthogonal factor rotation method for simplifying the columns of the factor matrix and was generally considered better than others (Hair Jr., 2009). Furthermore, factor structure refers to the overall fit of the two-factor solution (Cognitive-Affective and Somatic) reported in the BDI-II manual (Beck, Steer and Brown, 1996). We used factor loadings greater than 0.30 to indicate whether an item represents its factor, discarded communalities where variance was insufficient and required no fewer than three items for each factor (Handal and Lacey, 2017). The Kaiser-Meyer-Olkin (KMO) supports the use of factor analysis for the data, while Bartlett's test of sphericity is used to evaluate whether a correlation matrix is suitable for

factor analysis by testing the hypothesis. The good KMO correlation is above 0.60-0.70 (Netemeyer, Bearden and Sharma, 2003; Taherdoost, Sahibuddin and Jalaliyoon, 2022), and the significance p-value of Bartlett's test of sphericity is ($p < .05$) (Taherdoost, Sahibuddin and Jalaliyoon, 2022).

In order to evaluate the convergent validity of the Indonesian version of the BDI-II scale, we used composite reliability (CR) and average variance extracted (AVE). A CR > 0.70 should be considered acceptable, as should an AVE > 0.50 (El-Den et al., 2018; Muslih et al., 2021). Additionally, correlations between factors or constructs were used to test discriminant validity. We applied the Fornell and Larcker (FL) criterion and the Heterotrait-Monotrait (HTMT) correlation ratio to test discriminant validity. The Fornell-Larcker criterion suggests that each construct's AVE should have a greater value than the correlation with other latent constructs, whereas some authors suggest a threshold of < 0.85 or < 0.90 indicates good discriminant validity (HamidSami and Sidek, 2017). Furthermore, as per Plichta and Kelvin (2013), interim correlations between 0.3-0.7 and item-to-total correlations greater than 0.50 are acceptable (Nawi et al., 2020).

We used the CFA to evaluate the model fit of the Indonesia version BDI-II scale. The quality of fit was examined with the following indices: absolute fit measures are [χ^2/df , goodness-of-fit index (GFI), root mean square error of approximation (RMSEA), standard root-mean-square residual (SRMR)], additional fit measures are adjusted GFI (AGFI), Tucker–Lewis Index

(TLI), incremental fit index (IFI), normal fit index (NFI), comparative fit index (CFI), I and parsimonious fit measures [Parsimony NFI (PNFI), Parsimony GFI (PGFI)].

In detail, for model fit indices, the acceptable fit value for CFI, IFI, TLI, and NFI was > 0.90; $\chi^2/df < 3.0$; RMSEA and SRMR < 0.08; GFI and AGFI between 0.80 and 0.90; PNFI and PGFI > 0.50. Hair et al. (2009) recommended selecting GFI that meets at least one incremental and absolute fit measure.

Results

Demographic characteristics of samples

The sample for this study was local urban residents who are suffered from floods every year who were contacted through the district government and the East Java Provincial Political and National Unity Agency, and the Baleendah-Dayeuhkolot community health center in Bandung, West Java, Indonesia. A potential participant who withdrew from the start after stating his/her ability due to personal reasons was excluded from the further analysis process, leaving 107, of whom 67.29% were female. Ages ranged from 13 to 64 with a mean (SD of 32.19 (± 13.86)). The majority were single (77.57%) and 71.03% of respondents were unemployed. Table 1 shows about demographic characteristics of participants.

Reliability testing

Before conducting the reliability and validity test, we evaluated whether the data could proceed in the next step using the data normality test. The normality test for BDI-II showed a p-value of <0.001, which means that the

Table 2. Item-total correlation, Cronbach's alpha if item deleted, Cronbach's alpha, and validity

Item	Item-total correlation	Cronbach's alpha if item deleted	Validity
1. Sadness	0.47	0.89	Content validity: CVI: 0.97 S-CVI: 0.98
2. Pessimism	0.38	0.89	
3. Past failures	0.62	0.89	
4. Losing pleasure	0.55	0.89	Convergent validity AVE (overall): 0.31; CR: 0.90 AVE (factor 1): 0.32; CR: 0.83 AVE (factor 2): 0.30; CR: 0.81
5. Feeling of guilt	0.64	0.89	
6. Feeling of punishment	0.61	0.89	
7. Disliking myself	0.62	0.89	
8. Self-criticism	0.64	0.89	
9. Suicidal thoughts or desires	0.43	0.89	
10. Crying	0.29	0.90	
11. Agitation	0.57	0.89	
12. Losing interest	0.36	0.89	
13. Doubts	0.60	0.89	
14. Uselessness	0.62	0.89	Discriminant validity: HTMT: 0.86 FL: Square root of AVE: (0.56 and 0.54) < correlation of latent construct (0.72)
15. Loss of energy	0.63	0.89	
16. Changes in sleep patterns	0.39	0.89	
17. Offended	0.57	0.89	
18. Changes in appetite	0.53	0.89	
19. Difficulty in concentrating	0.72	0.89	
20. Fatigue	0.63	0.89	
21. Losing interest in sex	0.21	0.90	
Cronbach's alpha		0.89 (overall); 0.85 (factor 1); 0.81 (factor 2)	
Range of item-total correlation		0.21 to 0.72	

Note: CVI= Content Validity Index, S-CVI= Scale-level CVI, AVE= Average Variance Extracted, HTMT= Heterotrait-Monotrait, FL= Fornell & Larcker

Table 3. Factor pattern loadings of 21 items of EFA analysis of the BDI-II-Indonesian version

Item	Factor Loading	
	Factor 1 (Negative Cognitive-Attitude)	Factor 2 (Somatic Symptom)
1. Sadness	0.38	-
2. Pessimism	0.65	-
3. Past failures	0.74	-
4. Losing pleasure	0.61	-
5. Feeling of guilt	0.49	-
6. Feeling of punishment	0.72	-
7. Disliking myself	0.61	-
8. Self-criticism	0.30	-
9. Suicidal thoughts or desires	-	0.39
10. Crying	-	0.57
11. Agitation	-	0.60
12. Losing interest	0.54	-
13. Doubts	-	0.58
14. Uselessness	-	0.54
15. Loss of energy	-	0.58
16. Changes in sleep patterns	0.39	-
17. Offended	0.64	-
18. Changes in appetite	-	0.51
19. Difficulty in concentrating	-	0.51
20. Fatigue	-	0.68
21. Losing interest in sex	-	0.49
KMO and Bartlett's test of sphericity	0.857; p-value < 0.001	
Eigen value	7.53	1.51
% of variance	35.86%	7.19%
Cumulative % of the variance	35.86%	43.05%

Note: CVI= Content Validity Index, S-CVI= Scale-level CVI, AVE= Average Variance Extracted, HTMT= Heterotrait-Monotrait, FL= Fornell & Larcker

data did not have a normal distribution. We checked the Mahalanobis d-squared in CFA output and compared it with the χ^2 table ($\alpha= 0.001$) to find the outlier data. The χ^2 table value with df: 21 (observed variable) is 113.56 since the Mahalanobis d-squared values were < 113.56, thus no outlier data.

This study used Cronbach's alpha coefficient to assess the internal consistency of each factor and the general scale. The Cronbach's alpha coefficient of the overall IDN-BDI-II scale was 0.89. The Cronbach's alpha coefficient of each factor was 0.85 for the "Negative Cognitive-Attitude" 0.81 for "Somatic Symptom." Item-to-item correlation coefficients of the Indonesian BDI-II scale test showed an acceptable calculation for all subjects ranging from 0.19 (item-21 and item-2) to 0.67 (item-9 and item-15) (Appendix I). All subject's analyses generated results with item-to-total correlation ranging from 0.21 to 0.72 (Table 2).

Content Validity testing

We calculated the content validity through the I-CVI and S-CVI tests for the BDI-II scale, and the value of the I-CVI was 0.97, indicating that the item was relevant. On the other hand, the value of S-CVI was 0.98, indicating that the items had excellent content validity. Whereas CR tested the convergent validity, the result of the CR test overall was 0.90, and the values for "Negative Cognitive-Attitude" and "Somatic Symptom" were 0.83 and 0.81, respectively. The appropriate overall AVE values were 0.31 and values 0.32, and 0.30 for each

factor, respectively. The value of HTMT was 0.86, and the square root of AVE was 0.72, indicating good discriminant validity. All reliability and validity results tested are reported in Table 2.

Construct validity testing

We used SPSS Version 22 to evaluate the EFA, and AMOS version 23.0 software to perform CFA.

Exploratory factor analysis

Factor pattern loadings of 21 items are presented in Table 3. Generated factors based on an eigenvalue score greater than one indicates that the factor structure is considered to be maintained from these two factors (F1= 7.53 and F2=1.51). The KMO measure was 0.86, and Bartlett's test of sphericity was significant with a p-value < 0.001. The main components in this inventory were analyzed using varimax rotation and showed that the factor rotation value rotated on BDI-II for all items was greater than 0.30. Factor 1 (F1) covered symptoms belonging to the "Negative Cognitive-Attitude" domains (NCA), with symptoms such as past failures (0.74), feeling of punishment (0.72), pessimism (0.65), offended (0.64), disliking myself (0.61), losing pleasure (0.61), losing of interest (0.54), feeling guilty (0.49), changes in sleep pattern (0.39) sadness (0.38) and self-criticism (0.30) fulfilled standard factor loading on this factor. Factor 2 (F2) represented mainly "Somatic Symptoms" domains (SS), including fatigue (0.68), agitation (0.60), loss of energy (0.58), doubt (0.58),

crying (0.57), uselessness (0.54), changes in appetite (0.51), difficulty in concentrating (0.51) and losing interest in sex (0.49) and suicidal desires (0.39) (Table 3).

Confirmatory factor analysis

The model achieved good on all goodness-of-fit indicators, χ^2 ($\chi^2/df = 1.82$), p value <0.01 , CFI = 0.81, IFI = 0.82, TLI = 0.79, NFI = 0.67, RMSEA = 0.09, SRMR = 0.07, GFI = 0.8, AGFI = 0.72, PNFI = 0.72, PGFI = 0.63. Standardized regression weight and model fit in CFA for BDI-II are presented in Figure 1.

Discussions

This study explored the possibility of adapting the BDI-II for use in Indonesian populations in specific flood-suffered communities; several tests have been carried out using the backward-and-forward translation process, which can prove its applicability. This study found overall good internal consistency (Cronbach's alpha = 0.89) and the adaptability of BDI-II Indonesian version. Further, to fulfill the BDI-II level of depression in the original BDI-II version, the results show a more detailed picture of each level. The level of depression was at "normal ups and downs," with 59.8% ranging from the top position of all participants, and mild mood disturbance was in the second position, namely 27.1%. This finding is consistent with the results of Woody et al.'s (2017) study, which demonstrated excellent internal consistency and validity with Cronbach's alpha 0.94, on 51 women of the 2011 Binghamton floods sufferers.

Concerning current results, the EFA sample ($n = 107$) met the 5:1 subject-to-item ratio sample size criterion (Gorsuch, 1988). As with sample size estimation for EFA, our finding also met standard sample size recommendations for CFA, ranging between 100 and 200 subjects (Gagne and Hancock, 2006). Moreover, our KMO test results yielded 0.86, which indicates that the correlation in the BDI-II-Indonesian is above 0.60-0.70, which is considered sufficient to move toward factor analysis (Netemeyer, Bearden and Sharma, 2003; Taherdoost, Sahibuddin and Jalaliyoon, 2022).

At the same time, the participants showed borderline, moderately severe, and extreme, respectively. When referring to Gebrie's (2020) analysis of BDI-II, the depression level of the participants was mild depression. These results indicate that this scale can describe depression on specific demographic characteristics, namely: appropriate for the female gender as the majority of the sample in several countries: Mexico, Brazil, and Iran (Gebrie, 2020G). Also,

after validation, we found that major differences from two domains including "Negative- Cognitive Attitude" and "Somatic Symptom" original BDI-II with three domains.

This study used CFA to measure construct validity. At the same time, CR and AVE are used to measure convergent validity. Our study provides evidence that the BDI-II has sufficient validity. Several fit indices have been used to test data compatibility so that the found model structure can be tested. We kept items 10 and 21 with a loading factor of 0.26 and 0.23, respectively, because the overall goodness-of-fit model was still met. In addition, because crying (item 10) and losing interest in sex (item 21) are important indicators of depression. In general, all BDI-II scales factors indicated adequate factors. The 21st item on "Losing interest in sex" is closely related to cultural variables. Most Indonesians, especially conservative people, are still shy and even taboo in talking about "sex" (Hanifah, 2020). Therefore, it should be recommended to make the term "sex" more acceptable to all groups according to Indonesian culture and it is recommended to conduct more research on this area.

The inter-item correlations showed the result that the range of CR of factor 1 was 0.83 and CR factor 2 was 0.81 with an overall CR of 0.90, also indicating that the items were not redundant (no item had a high correlation (r greater than 0.7) each other) were satisfactory (Nawi et al., 2020).

This validation version of BDI-II in Bahasa Indonesia can be the right choice for particular specific population of flood survivors in the community setting. These results can generalize and answer the research questionnaire related to depression after disaster incidence by healthcare professionals and victims themselves, which is the strength of this study. We tried to implement a proper solution to minimize the gap of language barrier with the internationally published BDI-II version. However, this study has some limitations. The small number of samples affects the results in determining several loading factors. Thus, we recommend future research using a larger sample and multiple socio-cultural groups. Moreover, the public healthcare and mental health nurses' association and education can encourage using the BDI-II Indonesian version to screen depression severity properly for floods victims to test further reliability and validity. Nowadays digitalization of instrument is more common and easy to access. Therefore, after finishing all the psychometric testing within the Indonesia context, we would recommend to create a downloadable version of BDI-II.

Conclusions

This study illustrated the BDI-II Indonesian version with good validity and reliability and could be used to measure depression in the Indonesian community with Bahasa Indonesia, particularly in those who suffer from floods disasters. This version was validated mainly using Sundanese population. Further study needs to be validated using different social-cultural groups such as Minangkabau, Madurese, and Padangese to strengthen this psychometric property. Moreover, it should be tested with other provinces and states with multiple disaster incidences. Future study needs to be validated with different demographic factors such as age, gender and education as well as healthcare professionals. It is important to notice development of an internet version of BDI-II for easy accessibility in the current digital era.

Acknowledgment

The authors thank all the participants, Dr. Muhammad Muslih, M.S., for assisting and facilitating the peer discussion, Dean Faqih Ruhyanuddin for inspiring researchers to conduct this study, and all Mental Health Nursing experts in Indonesia Mental Health Nurses Association for sharing the practical experiences to support this translation.

Funding source

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflict of Interest

This study illustrated that the BDI-II Indonesian version had good validity and reliability; it indicated that the BDI-II Indonesian version was a good tool and could be used in measuring depression in the Indonesian community, particularly in those who suffer from floods disasters.

References

- Agusintadewi, N. (2019) 'Membangun Budaya Tanggap Bencana Berbasis Kearifan Lokal di Indonesia'. Conference: Seminar Nasional Arsitektur dan Tata Ruang (SAMARTA) 2019At: Denpasar, Bali
- Bei, B. *et al.* (2013) 'A prospective study of the impact of floods on the mental and physical health of older adults', *Aging & Mental Health*, 17(8), pp. 992-1002. <https://doi.org/https://doi.org/10.1080/13607863.2013.799119>
- Cacheda, F. *et al.* (2019) 'Early Detection of Depression: Social Network Analysis and Random Forest Techniques', *J Med Internet Res*, 21(6), pp. e12554. <https://doi.org/https://doi.org/10.2196/12554>
- El-Den, S. *et al.* (2018) 'The psychometric properties of depression screening tools in primary healthcare settings: A systematic review', *J Affect Disord*, 225, pp. 503-522.

- <https://doi.org/10.1016/j.jad.2017.08.060>
- Gagne, P. and Hancock, G. R. (2006) 'Measurement Model Quality, Sample Size, and Solution Propriety in Confirmatory Factor Models', *Multivariate Behavioral Research*, 41(1), pp. 65-83. https://doi.org/https://doi.org/10.1207/s15327906mbr4101_5
- Gan, Y. *et al.* (2014) 'Depression and the risk of coronary heart disease: a meta-analysis of prospective cohort studies', *BMC Psychiatry*, 14, pp. 371. <https://doi.org/https://doi.org/10.1186/s12888-014-0371-z>
- Gebrie, M. H. (2020) 'An analysis of BDI2nd edition (BDI-II)'.
 Ghassemzadeh, H. *et al.* (2005) 'Psychometric properties of a Persian-language version of the BDI- Second edition: BDI-II-PERSIAN', *Depression and anxiety*, 21, pp. 185-92. <https://doi.org/https://doi.org/10.1002/da.20070>
- Ginting, H. *et al.* (2013) 'Validating the Beck Depression Inventory-II in Indonesia's general population and coronary heart disease patients', *International Journal of Clinical and Health Psychology*, 13, pp. 235-242. [https://doi.org/https://doi.org/10.1016/S1697-2600\(13\)70028-0](https://doi.org/https://doi.org/10.1016/S1697-2600(13)70028-0)
- González, D. A. *et al.* (2015) 'Adaptation of the BDI-II in Mexico', *Salud Ment (Mex)*, 38(4), pp. 237-244. <https://doi.org/https://doi.org/10.17711/sm.0185-3325.2015.033>
- Gorsuch, R. L. (1988) 'Exploratory Factor Analysis', in Nesselroade, J.R. and Cattell, R.B. (eds.) *Handbook of Multivariate Experimental Psychology*. Boston, MA: Springer US, pp. 231-258. https://doi.org/10.1007/978-1-4613-0893-5_6
- Hanifah, A. (2020). Pendidikan Seks Di Kalangan Remaja (Kritik Ideologi Terhadap Teks Film Dua Garis Biru). <http://etheses.iainponorogo.ac.id/id/eprint/9057>
- Hair Jr., J. F. *et al.* (2009) 'Multivariate Data Analysis. 7th Edition, ' Prentice Hall, Upper Saddle River, 761.
- Hamid, M., Sami, W. and Sidek, M. (2017) 'Discriminant Validity Assessment: Use of Fornell & Larcker criterion versus HTMT Criterion', *Journal of Physics: Conference Series*, 890, pp. 012163. <https://doi.org/https://doi.org/10.1088/1742-6596/890/1/012163>
- Handal, P. J. and Lace, J. W. (2017) 'Differential Effects of Family Structure on Religion and Spirituality of Emerging Adult Males and Females', *J Relig Health*, 56(4), pp. 1361-1370. <https://doi.org/https://doi.org/10.1007/s10943-016-0343-3>
- Kadiyono, A. L., & Harding, D. (2017). Pengaruh nilai budaya sunda dalam upaya peningkatan kesejahteraan psikologis korban bencana tanah longsor. *Journal of Psychological Science and Profession*, 1(3), 26-35. <https://doi.org/10.24198/jpsp.v1i3.15232>
- Kellar, S. P. and Kelvin, E. A. 'Munro's statistical methods for health care research'.
- Mason, V., Andrews, H. and Upton, D. (2010) 'The psychological impact of exposure to floods', *Psychol Health Med*, 15(1), pp. 61-73. <https://doi.org/https://doi.org/10.1080/13548500903483478>
- Muslih, M. *et al.* (2021) 'Psychometric properties of the Internalized Stigma of Mental Illness (ISMI) Scale in Indonesia', *Psychiatric Rehabilitation Journal*, 44(2), pp. 166-175. <https://doi.org/https://doi.org/10.1037/prj0000462>
- Nawi, F. *et al.* (2020) 'A Review on The Internal Consistency of a Scale: The Empirical Example of The Influence of Human Capital Investment on Malcom Baldrige Quality Principles in Tvet Institutions', *Asian People Journal (APJ)*, 3, pp. 19-29. <https://doi.org/https://doi.org/10.37231/apj.2020.3.1.121>
- Netemeyer, R., Bearden, W. and Sharma, S. (2003) 'Scaling Procedures. Issues and Applications', *Sage*. <https://doi.org/https://doi.org/10.4135/9781412985772>
- Penninx, B. W. *et al.* (2013) 'Understanding the somatic consequences of depression: biological mechanisms and the role of depression symptom profile', *BMC Med*, 11, pp. 129. <https://doi.org/https://doi.org/10.1186/1741-7015-11-129>
- Rodrigues, I. B. *et al.* (2017) 'Development and validation of a new tool to measure the facilitators, barriers and preferences to exercise in people with osteoporosis', *BMC Musculoskeletal Disorders*, 18(1), pp. 540. <https://doi.org/10.1186/s12891-017-1914-5>
- Samson, C. M. S., Erwina, W., & Lusiana, E. (2021). Pengetahuan lokal dan strategi lokal Sunda dalam hadapi bencana. *Informatio: Journal of Library and Information Science*, 1(2), 181-194.

- <https://doi.org/10.24198/inf.v1i2.34268>
- Steger, M. F. and Kashdan, T. B. (2009) 'Depression and Everyday Social Activity, Belonging, and Well-Being', *J Couns Psychol*, 56(2), pp. 289-300. <https://doi.org/https://doi.org/10.1037/a0015416>
- Taherdoost, H., Sahibuddin, S. and Jalaliyoon, N. (2022) 'Exploratory factor analysis; concepts and theory', *Advances in applied and pure mathematics*, 27, pp. 375-382. <https://hal.science/hal-02557344/file/Exploratory%20Factor%20Analysis%3B%20Concepts%20and%20Theory.pdf>
- Telles, S., Singh, N. and Joshi, M. (2009) 'Risk of posttraumatic stress disorder and depression in survivors of the floods in Bihar, India', *Indian journal of medical sciences*, 63, pp. 330-4. <https://doi.org/https://doi.org/10.4103/0019-5359.55883>
- Wang, Y. P. and Gorenstein, C. (2013) 'Psychometric properties of the Beck Depression Inventory-II: a comprehensive review', *Braz J Psychiatry*, 35(4), pp. 416-31. <https://doi.org/https://doi.org/10.1590/1516-4446-2012-1048>
- Wind, T. R. et al. (2013) 'The Impact of Recurrent Disasters on Mental Health: A Study on Seasonal Floods in Northern India', *Prehospital and Disaster Medicine*, 28(3), pp. 279-285. <https://doi.org/https://doi.org/10.1017/S1049023X13000290>
- Woody, M. L. et al. (2017) 'Pupillary response to emotional stimuli as a risk factor for depressive symptoms following a natural disaster: The 2011 Binghamton flood', *Clin Psychol Sci*, 5(4), pp. 726-732. <https://doi.org/https://doi.org/10.1177/2167702617699932>
- Zamanzadeh, V. et al. (2015) 'Design and Implementation Content Validity Study: Development of an instrument for measuring Patient-Centered Communication', *J Caring Sci*, 4(2), pp. 165-78. <https://doi.org/10.15171/jcs.2015.017>
- Zhong, S. et al. (2018) 'The long-term physical and psychological health impacts of flooding: A systematic mapping', *Science of The Total Environment*, 626, pp. 165-194. <https://doi.org/https://doi.org/10.1016/j.scitotenv.2018.01.041>

How to cite this article: Ariani, T. A., Anna, A., Rahayu, H. T., Aini, N., Windarwati, H. D., Hernawaty, T., Mudiyansele, S. P.K., and Lin, M. F. (2023) 'Psychometric testing of the Indonesian version of Beck Depression Inventory-II among Indonesian floods survivors', *Jurnal Ners*, 18(3), pp. 264-273. doi: <http://dx.doi.org/10.20473/jn.v18i3.47313>

Appendix I. Item-to-item correlations for the BDI-II-Indonesian version

	item1	item2	item3	item4	item5	item6	item7	item8	item9	item10	item11	item12	item13	item14	item15	item16	item17	item18	item19	item20	item21
item1	1,000																				
item2	,106	1,000																			
item3	,357	,237	1,000																		
item4	,350	,338	,525	1,000																	
item5	,260	,265	,453	,357	1,000																
item6	,254	,452	,545	,385	,485	1,000															
item7	,355	,143	,617	,481	,428	,477	1,000														
item8	,355	,128	,428	,459	,470	,302	,414	1,000													
item9	,105	,156	,351	,242	,316	,129	,392	,301	1,000												
item10	,256	-,078	,086	,114	,263	,227	,137	,433	,002	1,000											
item11	,244	,253	,232	,318	,311	,376	,328	,456	,245	,228	1,000										
item12	,230	,332	,232	,171	,218	,386	,279	,190	,064	,094	,237	1,000									
item13	,226	,132	,508	,340	,379	,409	,555	,407	,484	,183	,401	,107	1,000								
item14	,300	,358	,322	,484	,539	,509	,430	,417	,296	,270	,467	,300	,426	1,000							
item15	,251	,287	,364	,415	,399	,358	,449	,469	,334	,121	,451	,195	,493	,395	1,000						
item16	,288	,221	,298	,121	,321	,268	,105	,264	,256	,061	,228	,356	,110	,109	,333	1,000					
item17	,284	,274	,630	,340	,469	,450	,425	,300	,361	,139	,236	,292	,427	,330	,223	,300	1,000				
item18	,371	,239	,286	,320	,287	,235	,358	,389	,265	,112	,395	,117	,319	,240	,492	,348	,345	1,000			
item19	,443	,392	,444	,441	,471	,464	,449	,517	,352	,135	,510	,224	,460	,486	,667	,371	,314	,528	1,000		
item20	,252	,317	,327	,233	,488	,300	,355	,498	,428	,254	,413	,190	,459	,375	,529	,307	,418	,459	,550	1,000	
item21	,022	-,019	,027	,040	,260	,115	,035	,097	,049	,205	,247	,126	,274	,095	,116	,124	,141	,100	,147	,260	1,000