Cross-cultural comparison of depressive symptoms on the Beck Depression Inventory-II, across six population samples

Marjo Seppänen, Tiina Lankila, Juha Auvinen, Jouko Miettunen, Raija Korpelainen and Markku Timonen

Background

The Beck Depression Inventory (BDI) and BDI-II (revised version) are some of the most widely used and comparable self-report scales for assessing the presence and severity of depressive symptoms in many countries. However, although the relative mean score of each symptom in different countries may vary, the cultural differences of BDI-II symptoms for each item have not been previously studied.

Aims

To examine the overall picture of the magnitude of the symptoms in the Finnish population, and compare the relative mean score of each symptom between all published population-based samples from different countries fulfilling the search criteria.

Method

We conducted a search for population-based studies reporting BDI-II item, using Scopus, PsycINFO and PubMed, and five population-based samples were identified. Relative average scores for each item of the scale were calculated for the Finnish population and five populations from other countries. Meta-regression methods were used to test the differences in the relative score of each symptom between each country separately, and results were then visually compared with spider charts.

Results

We found significant differences in several BDI-II item scores between countries: lower indecisiveness, higher changes in sleep pattern and higher irritability in Finland; higher loss of pleasure in Norway; higher loss of interest in the Dominic Republic; higher self-criticalness and feelings of punishment in Mexico; and higher sadness in Japan.

Conclusions

Based on the study fundings and including all currently published population-based samples with BDI-II scores, cultural differences in depressive symptoms should be considered when interpreting BDI-II item scores.

Keywords

Depressive disorders; epidemiology; rating scales; transcultural psychiatry; statistical methodology.

The Beck Depression Inventory (BDI)[1] and its revised version (BDI-II) are some of the most frequently used self-rating scales for measuring the severity of depressive symptoms.[2] The BDI was originally developed based on clinical experience and aimed to assess the varying intensity of depression.[1] It underwent two major revisions in 1978[3] and 1996,[4] as the BDI-IA and BDI-II, respectively. The BDI-II was modified to better recognise severe depression, possibly demanding hospital care.[4]

BDI-II in different populations

Reportedly, studies have shown that culture affects the way we express emotions,[5] often leading to misinterpretation of symptoms by clinicians.[6] Psychological research has mostly been conducted in Western countries and among Western study samples, which do not culturally represent global diversity.[7] Since the publication of the BDI-II, its psychometric properties have been studied on several different occasions, in different cultures and subgroups, although not always in representative samples.[8–10] Studies have shown that the reliability and validity of the BDI-II, measured by Cronbach's alpha, Spearman's rank correlation and Student's t-test, is good across different subgroups, including when used in different language versions.[11,12] On the contrary, different cut-off points have been suggested for the studied subgroup in some study settings, such as a cut-off score of ≥27 to differentiate veterans with mood disorder.[13] In addition, population-based studies on the psychometric properties and validity of the BDI-II included a description of the distribution of the item scores of the BDI-II.[12,14–17] However, there are few studies describing the distribution of symptoms in representative population-based samples.

The original BDI scale has been shown to be a valid screening measure for depression in Finland.[18] and a reliable tool for cross-cultural comparison in Europe, although some cross-cultural differences have been noticed, especially in the Spanish sample.[19] The Finnish translation of the BDI-II was published in 2004.[20] Since then, the BDI-II questionnaire has been used in clinical[21–23] and population-based studies,[24–26] but the individual items of the questionnaire have not been examined in the Finnish population since the translation. In addition, neither were the Finnish population symptom scores compared with the corresponding population scores from other countries.

Aim of the study

Although the BDI-II is a commonly used tool for measuring depressive symptoms worldwide, with some cultural differences being reported,[27,28] no item-by-item comparison of the BDI-II across different countries and population-based samples has been made. To improve the cross-cultural interpretation of the BDI-II, this population-based study described the symptoms measured on the BDI-II in the Finnish population, and compared the distribution of items between population-based samples from six different countries.
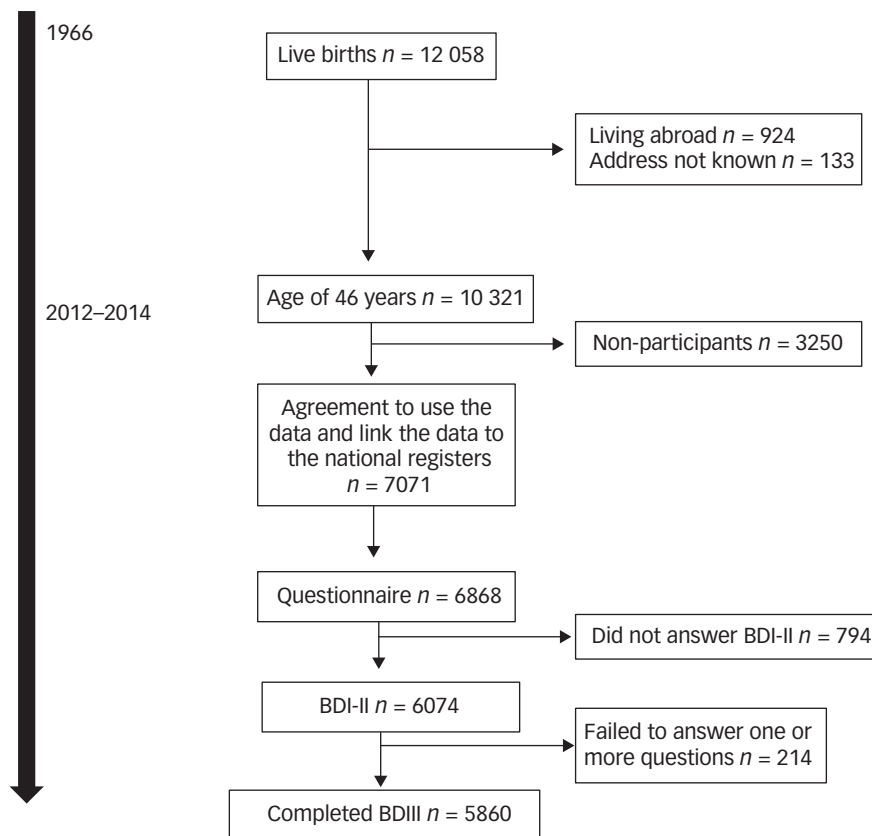
Fig. 1 Flow chart of the participants who completed the BDI-II questionnaire in the Northern Finland Birth Cohort 1966. BDI-II, Beck Depression Inventory-II.

Method

Finnish study population

The Finnish population was based on the Northern Finland Birth Cohort (NFBC) 1966,[29] which is a longitudinal research programme that originally included all of the mothers ($n = 12\,068$)[30] with children ($n = 12\,058$ live-born individuals) whose expected date of birth fell in the year 1966. The cohort members were monitored through interviews, postal questionnaires and clinical measurements from the prenatal period onward. The data from the most recent time point when the individuals were 46 years old ($n = 10\,321$ alive) were included in this study (see Fig. 1). Questionnaire data at 46 years were received from 6868 (67%) participants and clinical examination data were received from 5860 (57%) participants. The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008. All procedures involving human participants were approved by the Ethical Committee of the Northern Ostrobothnia Hospital District in Oulu, Finland (approval number 94/2011). Written informed consent was obtained from all participants. Personal identity information was encrypted and replaced with identification codes.

Selection of international populations

We searched SCOPUS, PsycINFO and PubMed for studies on the BDI-II. The search was conducted manually on 22 June 2020, using the following keywords: BDI-II, Beck Depression Inventory II, validation, population, psychometr*, adaptation and dimension.

The search strategy was developed in cooperation with a health science librarian using medical subject headings, and adapted for other databases by using free-word searches. The search strings were defined and validated by the first author (M.S.) together with the librarian. M.S. performed the initial search and screened the titles and abstracts of all of the articles identified by the search strings, to exclude irrelevant articles according to the eligibility criteria. Furthermore, we included original peer-reviewed journal articles, including mean values of BDI-II items based on population-based samples. The articles included in the analyses were required to meet the following criteria: used the BDI-II, reported the mean BDI-II value and mean and s.d. for each BDI-II item, and used population-based sampling. The search results were limited to human studies and English-language articles made available through open access or otherwise accessible within the University of Oulu. The search results were not limited by date. Additional articles were identified by searching the references in papers retrieved by the search strategy. The initial literature search retrieved 522 articles; however, after removing duplicate articles, screening of titles and abstracts, and reading of full-text articles, five articles were finally considered eligible for inclusion in this comparative study.

International populations

The eligible studies selected for the comparison were conducted in Norway ($n = 875$),[14] the Dominican Republic ($n = 954$),[12] Brazil ($n = 182$),[15] Mexico ($n = 205$)[16] and Japan ($n = 766$).[17] The selected studies are summarised in Table 1. The mean number of individuals included in the studies was 603 (s.d. 378), ranging from 182 to 954. Sample recruitment varied between studies. All of the study populations were adults or adolescents with a mean age of 40.6 years (range

Table 1 Summary table of studies included in the present comparison

Reference	Location	Sample selection	Sample size (women/men)	Mean age (s.d./range)	BDI-II translation	Cronbach's alpha	BDI-II total scores			
							0–13	14–19	20–28	29–63
NFBC, 1966[14]	Finland	Cohort study	5860 (3284/2576)	46.6 (0.6)	Finnish	0.92	4794 (82.0%)	594 (10.0%)	401 (4.1%)	71 (1.2%)
Aasen, 2001[14]	Norway	Randomly recruited by telephone register	875 (512/363)	44.5 (18.0–86.0)	Norwegian	0.91	711 (81.0%)	93 (10.6%)	52 (6.0%)	21 (2.4%)
García-Batista et al, 2018[12]	Dominican Republic	Convenience sample from the general public	954 (438/516)	31.6 (16.0)	Spanish	0.89	726 (76.1%)	102 (10.7%)	90 (9.4%)	36 (3.8%)
Gomes-Oliveira et al, 2012[15]	Brazil	Community-dwelling adults, population-based household survey	182 (102/80)	41.0 (10.8)	Brazilian Portuguese	0.93	NA	NA	NA	NA
González et al, 2015[16]	Mexico	Community sample	205 (117/88)	29.9 (16.0–70.0)	Spanish	0.87	160 (78%)	27 (13%)	10 (5%)	8 (4%)
Kojima et al, 2002[17]	Japan	General population recruited from annual communal health check-up visit	766 (303/463)	50.0 (8.6)	Japanese	0.87	618 (80.7%)	99 (12.9%)	37 (4.8%)	12 (1.6%)

BDI-II, Beck Depression Inventory-II; NFBC, Northern Finland Birth Cohort.

29.9–50.0 years). The studies were conducted in 2001, 2002, 2012, 2015 and 2018.

Instrument

Depression was evaluated with the BDI-II questionnaire, which consists of 21 questions and answers, each scoring from 0 to 3 points. The total score of BDI-II can be calculated (0–63), and a cut-off of 0–13 points indicates minimal depressive symptoms, 14–19 points indicates mild depressive symptoms, 20–28 points indicates moderate depressive symptoms and 29–63 points indicates severe depressive symptoms.[4] All study populations used a translated version of the questionnaire, and samples had adequate internal consistency as measured by Cronbach's alpha, ranging from 0.87 to 0.93. We present the detailed information regarding cut-off distributions and measures for internal consistency in Table 1.

Statistical analysis

We present the original BDI-II item means for each study in Supplementary Table 1 available at https://doi.org/10.1192/bjo.2022.13. As the samples differed by total BDI-II scores and we wanted to compare the relative importance of each item, we calculated the relative means for the items by dividing the mean value of each symptom with the total mean, and then divided by the number of BDI-II items (item mean/(total mean/21)). A relative mean <1 in an item can be interpreted so that the symptom in question has a mean score below an average item mean in the sample in question; if the mean is >1, the mean score is above an average item mean. Based on these relative mean scores, we first used random-effects meta-analysis to pool the six different populations, and then used meta-regression to statistically compare the results of these populations. This was done by comparing each sample to the average of the other samples. Visual examination of relative mean scores and results of meta-regression was done by using spider charts. The results are expressed as *P*-values and 95% confidence intervals. Statistical significance was determined as $P < 0.05$. Analyses were performed with IBM SPSS Statistics version 25 for Windows, and meta-analysis was conducted with Stata version 16 for Windows.

Results

Finnish population

Overall, 6074 participants completed the BDI-II questionnaire either on the day of the health examination or at home after receiving the questionnaire by mail; 214 participants failed to answer 1–14 questions. All of the those with missing values were excluded. The final study sample consisted of 5860 participants, of whom 2576 (46.2%) were men and 3284 (53.8%) were women. Most of the participants had good or excellent self-rated health (*n* = 3755, 64.1%), and 1577 (26.9%) had higher education. Of those who failed to complete the BDI-II questionnaire, 92 (43.0%) were men and 122 (57.0%) were women. The three items most likely to be missed were 'loss of energy' (*n* = 140, 65.4% missing values), 'tiredness or fatigue' (*n* = 139, 65.0%) and 'loss of interest in sex' (*n* = 139, 65.0%). Women were most likely not to respond to the question regarding loss of energy, whereas men were most likely not to respond to the question regarding loss of interest in sex. The average total of BDI-II points was 5.55 (s.d. 0.08), with a minimum value of 0 and a maximum value of 55. Ten per cent (*n* = 594) of the final study population scored at least 14 points (showing mild depressive symptoms) in the BDI-II questionnaire, 4.1% (*n* = 401) scored at least 20 points (showing moderate depressive symptoms) and 1.2% (*n* = 71) scored at least 29 points (showing
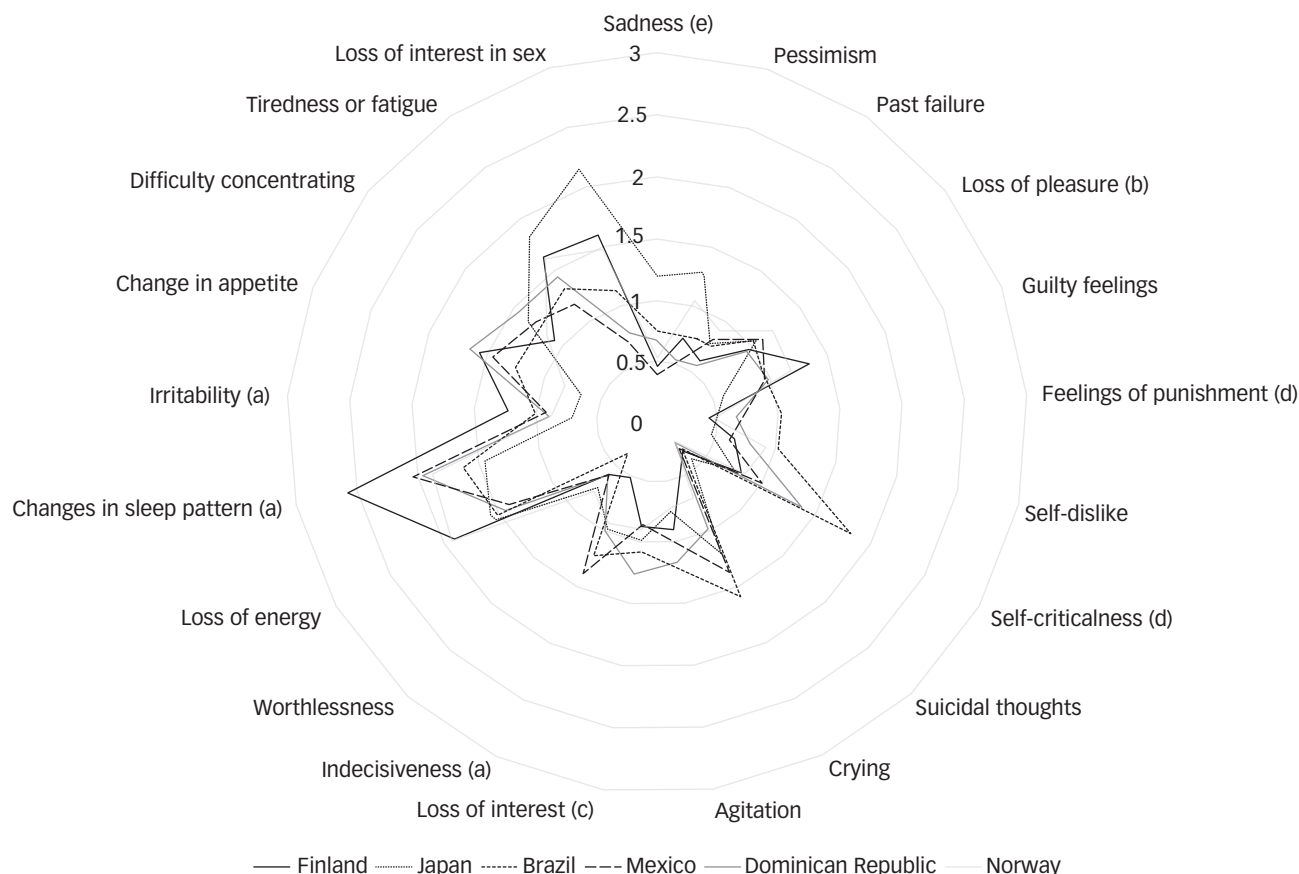
Fig. 2 Spider chart showing relative means of BDI-II items in different culture populations. Significant difference (*P* < 0.05) between all other countries and Finland (a), Norway (b), Dominican Republic (c), Mexico (d) and Japan (e). BDI-II, Beck Depression Inventory-II.

severe depression symptoms). Most of the individuals with depression were women (66.2%). The mean values of BDI-II in the non-depressed and depressed groups were 3.89 (s.d. 3.55) and 20.28 (s.d. 6.68), respectively. Means and s.d. for BDI-II items in Finnish population are shown in Supplementary Table 1.

Comparison of study populations from different countries

The mean scores of each symptom of BDI-II in different cultural populations are shown in Supplementary Table 1. The item 'changes in sleep pattern' (range 0.68–0.95) scored the highest in Finland, the Dominican Republic, Brazil and Mexico; 'loss of energy' (0.74) scored the highest in Norway; and 'loss of interest in sex' (0.91) scored the highest in Japan. 'Suicidal thoughts' (range 0.08–0.17) scored the lowest in every population. The item that scored the second lowest was 'feelings of punishment' in Finland and Norway (Finland: 0.11, Norway: 0.18), 'worthlessness' in Brazil and Mexico (Brazil: 0.27, Mexico: 0.16), 'pessimism' in the Dominican Republic (0.25) and 'self-dislike' in Japan (0.19).

The symptom items according to the BDI-II in six different cultural populations and the statistically significant differences between the populations are presented in Fig. 2. The Finnish population scored significantly lower in 'indecisiveness' (*P* = 0.034) and significantly higher in 'changes in sleep pattern' (*P* = 0.039) and 'irritability' (*P* = 0.019) than other populations. Compared with other populations, Norway scored significantly higher in 'loss of pleasure' (*P* = 0.033), the Dominican Republic scored significantly higher in 'loss of interest' (*P* = 0.009), Mexico scored significantly higher in 'self-criticalness' (*P* = 0.049) and 'feelings of punishment' (*P* = 0.048), and

Japan scored significantly higher in 'sadness' (*P* = 0.013). The detailed results are shown in Table 2.

Discussion

We compared the BDI-II item scores in all currently available population-based samples from six different countries, and found significant differences in several item scores between them. In the Finnish population, the item 'indecisiveness' scored lower and items 'changes in sleep pattern' and 'irritability' scored higher than in other populations. The Japanese population had a significantly higher 'sadness' score, the Norwegian population had a higher 'loss of pleasure' score, the Mexican population had higher 'self-criticalness' and 'feelings of punishment' scores, and the Dominican Republic population reported significantly higher 'loss of interest' score than samples from the other countries. Thus, the findings of this study can increase awareness of the cultural differences in depressive symptoms and enable effective interpretation of the BDI-II item scores.

To the best of our knowledge, this is the first study that conducted an item-by-item comparison of the BDI-II across different countries, with population-based samples. Consequently, the possibility of comparing our results with earlier studies is limited. When considering ethnicity and cultural background, no significant correlation between ethnicity (categorised as White versus other) and BDI-II total score was previously found.[4,31] Some measurement invariance defined by ethnicity has been reported across groups.[27,28] In addition, cultural differences in symptom reporting were revealed among Chinese-heritage and European-heritage

Table 2 Results of meta-analysis including relative mean scores for each symptom in each country and meta-regression comparing the differences in the relative score of each symptom between each country separately

	Relative mean scores		Meta-regression	
	Mean[a]	95% CI	Coefficient	95% CI
Sadness				
Finland	0.45	0.42–0.49	−0.24	−1.16 to 0.67
Norway	0.52	0.44–0.60	−0.17	−1.12 to 0.79
Dominican Republic	0.66	0.57–0.76	0.01	−0.98 to 1.00
Brasilia	0.75	0.54–0.95	0.10	−0.90 to 1.11
Mexico	0.39	0.26–0.51	−0.32	−1.21 to 0.56
Japan	1.18	1.10–1.28	0.64	**0.22–1.06**
Pessimism				
Finland	0.72	0.67–0.76	−0.11	−1.09 to 0.87
Norway	1.03	0.94–1.13	0.27	−0.65 to 1.19
Dominican Republic	0.54	0.45–0.62	−0.33	−1.21 to 0.55
Brasilia	0.75	0.54–0.95	−0.08	−1.09 to 0.94
Mexico	0.54	0.37–0.70	−0.33	−1.23 to 0.58
Japan	1.27	1.15–1.40	0.56	−0.77 to 1.19
Past failure				
Finland	0.61	0.56–0.65	−0.15	−0.55 to 0.26
Norway	0.91	0.80–1.02	0.23	−0.10 to 0.56
Dominican Republic	0.56	0.48–0.64	−0.20	−0.58 to 0.17
Brasilia	0.77	0.53–1.00	0.05	−0.49 to 0.59
Mexico	0.81	0.60–1.03	0.10	−0.41 to 0.62
Japan	0.76	0.66–0.85	0.39	−0.43 to 0.50
Loss of pleasure				
Finland	0.95	0.90–0.99	−0.10	−0.40 to 0.20
Norway	1.20	1.09–1.30	0.22	**0.03–0.41**
Dominican Republic	0.92	0.82–1.03	−0.12	−0.45 to 0.20
Brasilia	1.09	0.83–1.34	0.07	−0.40 to 0.54
Mexico	1.01	0.81–1.20	−0.02	−0.44 to 0.39
Japan	1.04	0.95–1.13	0.17	−0.37 to 0.37
Guilty feelings				
Finland	1.32	1.27–1.38	0.41	−0.26 to 1.08
Norway	1.16	1.07–1.26	0.22	−0.63 to 1.06
Dominican Republic	0.96	0.87–1.06	−0.02	−0.91 to 0.87
Brasilia	0.94	0.72–1.15	−0.06	−0.98 to 0.86
Mexico	0.94	0.79–1.09	−0.05	−0.95 to 0.85
Japan	0.57	0.49–0.64	−0.51	−1.04 to 0.02
Feelings of punishment				
Finland	0.42	0.38–0.45	−0.21	−0.76 to 0.34
Norway	0.47	0.37–0.56	−0.15	−0.77 to 0.47
Dominican Republic	0.64	0.54–0.75	0.06	−0.60 to 0.72
Brasilia	0.68	0.42–0.94	0.11	−0.63 to 0.84
Mexico	1.01	0.73–1.28	0.49	**0.01–0.97**
Japan	0.50	0.41–0.58	−0.12	−0.75 to 0.52
Self-dislike				
Finland	0.64	0.60–0.68	−0.09	−0.74 to 0.56
Norway	0.91	0.79–1.02	0.23	−0.36 to 0.82
Dominican Republic	0.77	0.67–0.87	0.06	−0.61 to 0.73
Brasilia	0.60	0.38–0.81	−0.14	−0.84 to 0.56
Mexico	1.01	0.76–1.25	0.33	−0.29 to 0.95
Japan	0.45	0.37–0.53	−0.32	−0.78 to 0.13
Self-criticalness				
Finland	0.80	0.75–0.84	−0.34	−1.69 to 1.00
Norway	0.83	0.72–0.93	−0.30	−1.67 to 1.07
Dominican Republic	1.37	1.26–1.48	0.35	−0.99 to 1.70
Brasilia	1.00	0.75–1.25	−0.09	−1.56 to 1.37
Mexico	1.82	1.56–2.07	0.88	**0.01–1.76**
Japan	0.71	0.61–0.81	−0.45	−1.73 to 0.84
Suicidal thoughts				
Finland	0.30	0.28–0.33	0.04	−0.21 to 0.29

(Continued)

Table 2 (Continued)

	Relative mean scores		Meta-regression	
	Mean[a]	95% CI	Coefficient	95% CI
Norway	0.21	0.16–0.26	−0.08	−0.31 to 0.15
Dominican Republic	0.21	0.16–0.27	−0.07	−0.31 to 0.17
Brasilia	0.21	0.10–0.33	−0.07	−0.35 to 0.21
Mexico	0.28	0.17–0.39	0.01	−0.28–0.29
Japan	0.40	0.33–0.47	1.55	−0.01 to 0.32
Crying				
Finland	0.45	0.41–0.50	−0.67	−1.67 to 0.33
Norway	0.67	0.56–0.79	−0.41	−1.69 to 0.87
Dominican Republic	0.96	0.84–1.09	−0.06	−1.46 to 1.34
Brasilia	1.36	1.04–1.68	0.41	−0.94 to 1.76
Mexico	1.56	1.23–1.89	0.65	−0.53 to 1.83
Japan	1.18	1.08–1.28	0.20	−1.17 to 1.57
Agitation				
Finland	0.87	0.83–0.91	−0.08	−0.79 to 0.63
Norway	0.70	0.61–0.79	−0.28	−0.86 to 0.31
Dominican Republic	1.14	1.02–1.26	0.25	−0.37 to 0.88
Brasilia	1.02	0.75–1.29	0.10	−0.69 to 0.89
Mexico	1.24	0.98–1.50	0.36	−0.30 to 1.02
Japan	0.73	0.63–0.83	−0.24	−0.88 to 0.39
Loss of interest				
Finland	0.87	0.83–0.92	−0.13	−0.60 to 0.34
Norway	0.88	0.78–0.98	−0.12	−0.61 to 0.38
Dominican Republic	1.24	1.13–1.35	0.34	**0.14–0.54**
Brasilia	0.83	0.60–1.07	−0.17	−0.73 to 0.40
Mexico	1.05	0.82–1.28	0.08	−0.49 to 0.66
Japan	0.97	0.87–1.06	−0.01	−0.53 to 0.51
Indecisiveness				
Finland	0.49	0.45–0.53	−0.54	**−1.01 to −0.07**
Norway	0.91	0.81–1.00	−0.07	−1.05 to 0.91
Dominican Republic	0.96	0.84–1.09	0.004	−0.98 to 0.99
Brasilia	1.36	1.07–1.65	0.47	−0.38 to 1.33
Mexico	1.20	0.96–1.44	0.28	−0.67 to 1.24
Japan	0.94	0.84–1.04	−0.02	−1.00 to 0.96
Worthlessness				
Finland	0.57	0.53–0.61	−0.03	−0.49 to 0.43
Norway	0.75	0.65–0.85	0.19	−0.19 to 0.58
Dominican Republic	0.58	0.48–0.68	0.01	−0.49 to 0.46
Brasilia	0.57	0.38–0.77	−0.02	−0.55 to 0.51
Mexico	0.34	0.21–0.48	−0.30	−0.60 to 0.01
Japan	0.71	0.61–0.81	0.14	−0.29 to 0.57
Loss of energy				
Finland	1.89	1.84–1.95	0.33	−0.32 to 0.99
Norway	1.91	1.81–2.02	0.36	−0.30 to 1.01
Dominican Republic	1.41	1.31–1.52	−0.24	−0.98 to 0.49
Brasilia	1.49	1.25–1.72	−0.15	−0.98 to 0.69
Mexico	1.37	1.16–1.57	−0.29	−1.05 to 0.46
Japan	1.56	1.46–1.65	0.07	−0.88 to 0.74
Changes in sleep pattern				
Finland	2.57	2.50–2.64	0.80	**0.06–1.52**
Norway	1.86	1.72–2.00	−0.07	−1.41 to 1.28
Dominican Republic	1.95	1.83–2.07	0.04	−1.31 to 1.39
Brasilia	1.62	1.34–1.90	−0.36	−1.66 to 0.95
Mexico	2.03	1.81–2.25	0.14	−1.22 to 1.49
Japan	1.44	1.33–1.55	−0.58	−1.65 to 0.49
Irritability				
Finland	1.21	1.16–1.26	0.38	**0.10–0.66**
Norway	0.85	0.75–0.95	−0.07	−0.70 to 0.56
Dominican Republic	0.88	0.77–0.98	−0.04	−0.67 to 0.59
Brasilia	0.98	0.74–1.22	0.78	−0.61 to 0.77
Mexico	0.86	0.67–1.04	−0.06	−0.72 to 0.59

(Continued)

Table 2 (Continued)

	Relative mean scores		Meta-regression	
	Mean[a]	95% CI	Coefficient	95% CI
Japan	0.68	0.59–0.78	−0.28	−0.78 to 0.22
Change in appetite				
Finland	1.55	1.50–1.61	0.40	−0.86 to 1.67
Norway	0.80	0.70–0.90	−0.50	−1.71 to 0.71
Dominican Republic	1.63	1.50–1.76	0.49	−0.72 to 1.71
Brasilia	1.23	0.97–1.50	0.02	−1.41 to 1.44
Mexico	1.45	1.21–1.70	0.28	−1.09 to 1.65
Japan	0.66	0.58–0.75	−0.67	−1.70 to 0.36
Difficulty concentrating				
Finland	1.06	1.01–1.11	−0.23	−0.58 to 0.11
Norway	1.14	1.04–1.24	−0.13	−0.61 to 0.35
Dominican Republic	1.44	1.31–1.56	0.24	−0.16 to 0.64
Brasilia	1.23	0.99–1.48	−0.01	−0.59 to 0.57
Mexico	1.28	1.07–1.50	0.05	−0.51 to 0.60
Japan	1.35	1.25–1.44	0.12	−0.34 to 0.59
Tiredness or fatigue				
Finland	1.63	1.57–1.68	0.14	−0.62 to 0.90
Norway	1.60	1.50–1.71	0.11	−0.67 to 0.89
Dominican Republic	1.44	1.32–1.55	−0.09	−0.88 to 0.69
Brasilia	1.32	1.07–1.57	−0.23	−1.03 to 0.58
Mexico	1.18	0.99–1.36	−0.41	−1.02 to 0.21
Japan	1.84	1.75–1.94	0.39	−0.16 to 0.94
Loss of interest in sex				
Finland	1.59	1.52–1.65	0.35	−1.47 to 2.17
Norway	1.47	1.34–1.61	0.22	−1.66 to 2.08
Dominican Republic	0.77	0.66–0.88	−0.63	−2.30 to 1.03
Brasilia	1.13	0.85–1.40	−0.20	−2.10 to 1.69
Mexico	0.66	0.47–0.86	−0.76	−2.35 to 0.82
Japan	2.15	2.00–2.30	1.02	−0.25 to 2.28

Figures marked in bold indicate statistically significant results.
a. Relative means score calculated by dividing the mean value of each symptom with the total mean and further divided by the number of Beck Depression Inventory-II items (item mean/(total mean/21)).

undergraduates in North America, as Chinese-heritage students scored higher on cognitive symptoms of depression.[27] In our study, Asia was only represented by Japan, which scored higher in the 'sadness' item, which is part of the general depression symptoms factor in the model presented by Ward.[32]

The original version of the BDI has previously been compared between European countries, and although it has been suggested to be a good tool for cross-cultural comparison, some differences in the relative weight of BDI items have been found between countries, especially regarding the Spanish sample, which placed greater importance on items 'sadness', 'pessimism' and 'self-accusation', and less importance on items 'guilty feelings', 'indecisiveness' and 'loss of libido'. In addition, a Finnish sample placed greater importance on items 'social withdraw' and 'body image', and a British sample placed greater importance on items 'fatigability' and 'weight loss'.[19] In our study, the Dominican Republic and Mexico samples used the Spanish-language version of the BDI-II, and found similar results regarding the Mexican sample scoring higher than our other samples on the items 'feelings of punishment' and 'self-criticalness'.

The probability of answering certain items of the BDI-II with low or high points may have been influenced by underlying cultural or language-version issues. This study was not focused on different language versions, but compared studies that used a translation of the BDI-II. Our study supports the previous recommendation that cultural context should be taken into better account when assessing issues of clinical psychology.[33] We can only speculate on the clinical settings, but taking account that patients likely have culturally diverse background, the cultural background and the potential role of acculturation should also be acknowledged when interpreting the results of questionnaire and when choosing the language version the patient will answer.

The major strength of this study is that this is the first international, population-based, item-by-item comparison made for BDI-II. However, this study has some limitations, including the differences between the populations we compared. Not all of the study participants were recruited through random sampling. The study populations also differed by age, although most of them were focused on adults. We found differences in item scores from the two younger samples (Dominican Republic and Mexico), which could also be associated with the lower age of the participants. All of the studies used different translations instead of the original BDI-II, which may partly explain the differences between the study populations, as translations might slightly alter the meaning of items. It is also important to note that we only had data from six countries, so it is not possible to draw strong conclusions about cultural differences, and comparison based on larger data is needed in the future. Thus, there is a need to examine the underlying factors explaining the cultural differences in future studies, as cultural background might affect how people answer the questionnaire and thus impede the results.

In conclusion, we found distinct differences in BDI-II item scores distribution between different countries. Thus, the possible cultural or language differences should be considered when interpreting the BDI-II questionnaire scores or comparing the BDI-II scores cross-culturally, as failure to do so may impede findings.

Marjo Seppänen (iD), Geography Research Unit, University of Oulu, Finland; Center for Life Course Health Research, University of Oulu, Finland; and Department of Sports and Exercise Medicine, Oulu Deaconess Institute Foundation, Finland; **Tiina Lankila**, Geography Research Unit, University of Oulu, Finland; and Department of Sports and Exercise Medicine, Oulu Deaconess Institute Foundation, Finland; **Juha Auvinen**, Center for Life Course Health Research, University of Oulu, Finland; and Medical Research Center, Oulu University Hospital and University of Oulu, Finland; **Jouko Miettunen** (iD), Center for Life Course Health Research, University of Oulu, Finland; and Medical Research Center, Oulu University Hospital and University of Oulu, Finland; **Raija Korpelainen**, Center for Life Course Health Research, University of Oulu, Finland; Department of Sports and Exercise Medicine, Oulu Deaconess Institute Foundation, Finland; and Medical Research Center, Oulu University Hospital and University of Oulu, Finland; **Markku Timonen**, Center for Life Course Health Research, University of Oulu, Finland; and Unit of General Practice, Oulu University Hospital, Finland

Correspondence: Marjo Seppänen. Email: marjo.seppanen@oulu.fi

First received 26 Aug 2021, final revision 7 Jan 2022, accepted 17 Jan 2022

Supplementary material

Data availability

Acknowledgements

Author contributions

References

1 Beck AT, Ward CH, Mendelson M, Mock J, Erbaugh J. An inventory for measuring depression. *Arch Gen Psychiatry* 1961; **4**: 561–71.

2 Richter P, Werner J, Heerlein A, Kraus A, Sauer H. On the validity of the Beck Depression Inventory. A review. *Psychopathology* 1998; **31**(3): 160–8.

3 Beck AT, Rush AJ, Shaw BF, Emery G. *Cognitive Therapy of Depression*. Guilford Press, 1979.

4 Beck AT, Steer RA, Brown G. *Manual for Beck Depression Inventory-II*. Psychological Corporation, 1996.

5 Markus HR, Kitayama S. Cultures and selves: a cycle of mutual constitution. *Perspect Psychol Sci* 2010; **5**(4): 420–30.

6 Kirmayer LJ. Cultural variations in the clinical presentation of depression and anxiety: implications for diagnosis and treatment. *J Clin Psychiatry* 2001; **62** (suppl 13): 22–30.

7 Thalmayer AG, Toscanelli C, Arnett JJ. The neglected 95% revisited: is American psychology becoming less American? *Am Psychol* 2021; **76**(1): 116–29.

8 Arnarson PÖ, Ólason DP, Smári J, Sigurdsson JF. The Beck Depression Inventory Second Edition (BDI-II): psychometric properties in Icelandic student and patient populations. *Nord J Psychiatry* 2008; **62**(5): 360–5.

9 Naja S, Al-Kubaisi N, Chehab M, Al-Dahshan A, Abuhashem N, Bougmiza I. Psychometric properties of the Arabic version of EPDS and BDI-II as a screening tool for antenatal depression: evidence from Qatar. *BMJ Open* 2019; **9**(9): e030365.

10 Harris CA, D'Eon JL. Psychometric properties of the Beck Depression Inventory-Second Edition (BDI-II) in individuals with chronic pain. *Pain* 2008; **137**(3): 609–22.

11 Toledano-Toledano F, Contreras-Valdez JA. Validity and reliability of the Beck Depression Inventory II (BDI-II) in family caregivers of children with chronic diseases. *PLoS One* 2018; **13**(11): e0206917.

12 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**(6): e0199750.

13 Reis DJ, Namekata MS, Oehlert ME, King N. A preliminary review of the Beck Depression Inventory-II (BDI-II) in veterans: are new norms and cut scores needed? *Psychol Serv* 2020; **17**(3): 363–71.

14 Aasen H. *An empirical investigation of depression symptoms: norms, psychometric characteristics and factor structure of the Beck Depression Inventory-II*. *Master's thesis*, The University of Bergen, 2001.

15 Gomes-Oliveira MH, Gorenstein C, Neto FL, Andrade LH, Wang YP. Validation of the Brazilian Portuguese version of the Beck Depression inventory-II in a community sample. *Rev Bras Psiquiatr* 2012; **34**(4): 389–94.

16 González DA, Rodríguez AR, Reyes-Lagunes I. Adaptation of the BDI-II in Mexico. *Salud Ment* 2015; **38**(4): 237–44.

17 Kojima M, Furukawa TA, Takahashi H, Kawai M, Nagaya T, Tokudome S. Cross-cultural validation of the Beck Depression Inventory-II in Japan. *Psychiatry Res* 2002; **110**(3): 291–9.

18 Nuevo R, Lehtinen V, Reyna-Liberato PM, Ayuso-Mateos JL. Usefulness of the Beck Depression Inventory as a screening method for depression among the general population of Finland. *Scand J Public Health* 2009; **37**(1): 28–34.

19 Nuevo R, Dunn G, Dowrick C, Vázquez-Barquero JL, Casey P, Dalgard OS, et al. Cross-cultural equivalence of the Beck Depression Inventory: a five-country analysis from the ODIN study. *J Affect Disord* 2008; **114**(1): 156–62.

20 Psykologien Kustannus Oy. *BDI-II Käsikirja [Beck Depression Inventory-II]*. Psykologien Kustannus Oy, 2005.

21 Granö N, Salmijärvi L, Karjalainen M, Kallionpää S, Roine M, Taylor P. Early signs of worry: psychosis risk symptom visual distortions are independently associated with suicidal ideation. *Psychiatry Res* 2015; **225**(3): 263–7.

22 Wäljas M, Iverson GL, Hartikainen KM, Liimatainen S, Dastidar P, Soimakallio S, et al. Reliability, validity and clinical usefulness of the BNI fatigue scale in mild traumatic brain injury. *Brain Inj* 2012; **26**(7–8): 972–8.

23 Weizmann-Henelius G. Violent female perpetrators in Finland. *Nord Psychol* 2006; **58**(4): 280–97.

24 Hintsa T, Wesolowska K, Elovainio M, Strelau J, Pulkki-Råback L, Keltikangas-Järvinen L. Associations of temporal and energetic characteristics of behavior with depressive symptoms: a population-based longitudinal study within Strelau's regulative theory of temperament. *J Affect Disord* 2016; **197**: 196–204.

25 Perkkiö Y, Jokelainen J, Auvinen J, Eskola P, Saltevo J, Keinänen-Kiukaanniemi S, et al. Glucose status and depressive symptoms: a cohort study of elderly people in Northwest Finland. *Scand J Prim Health Care* 2019; **37**(2): 242–8.

26 Poutanen O, Koivisto A-M, Salokangas RKR. Applicability of the DEPS Depression Scale: assessing format and individual items in subgroups of patients. *Nord J Psychiatry* 2010; **64**(6): 384–90.

27 Dere J, Watters CA, Yu SC-M, Michael Bagby R, Ryder AG, Harkness KL. Cross-cultural examination of measurement invariance of the Beck Depression Inventory-II. *Psychol Assess* 2015; **27**(1): 68–81.

28 Whisman MA, Judd CM, Whiteford NT, Gelhorn HL. Measurement invariance of the Beck Depression Inventory-Second Edition (BDI-II) across gender, race, and ethnicity in college students. *Assessment* 2013; **20**(4): 419–28.

29 University of Oulu. *Northern Finland Birth Cohort 1966*. University of Oulu, 1966 (http://urn.fi/urn:nbn:fi:att:bc1e5408-980e-4a62-b899-43bec3755243).

30 Rantakallio P. The longitudinal study of the Northern Finland birth cohort of 1966. *Paediatr Perinat Epidemiol* 1988; **2**(1): 59–88.

31 Steer RA, Ball R, Ranieri WF, Beck AT. Dimensions of the Beck Depression Inventory-II in clinically depressed outpatients. *J Clin Psychol* 1999; **55**(1): 117–28.

32 Ward LC. Comparison of factor structure models for the Beck Depression Inventory–II. *Psychol Assess* 2006; **18**(1): 81–8.

33 Ryder AG, Ban LM, Chentsova-Dutton YE. Towards a cultural-clinical psychology. *Soc Personal Psychol Compass* 2011; **5**(12): 960–75.