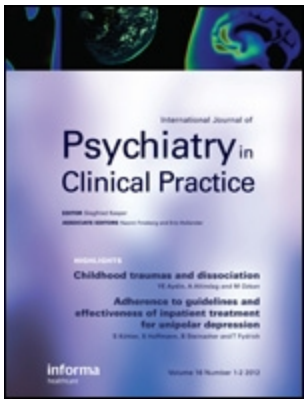


Beck Depression Inventory, Second Edition, Czech Version: Demographic Correlates, Factor Structure, and Comparison with Foreign Data

Journal:	*International Journal of Psychiatry in Clinical Practice*
Manuscript ID	MPCP-2019-0017.R2
Manuscript Type:	Original Article
Date Submitted by the Author:	n/a
Complete List of Authors:	Čihařová, Markéta; Vrije Universiteit Amsterdam, Section Clinical Psychology, Faculty of Behavioural and Movement Sciences; Charles University in Prague 1st Faculty of Medicine, and General University Hospital in Prague, Department of Neurology and Centre of Clinical Neuroscience Cígler, Hynek; Masarykova Univerzita Fakulta socialnich studii, Department of Psychology & Institute for Research on Children, Youth, and Family, Faculty of Social Studies Dostálová, Veronika; Charles University in Prague 1st Faculty of Medicine, and General University Hospital in Prague, Department of Neurology and Centre of Clinical Neuroscience Šivicová, Gabriela; Charles University in Prague, Faculty of Arts, Department of Psychology Bezdicek, Ondrej; Charles University in Prague 1st Faculty of Medicine, and General University Hospital in Prague, Department of Neurology and Centre of Clinical Neuroscience; National Institute of Mental Health
Keywords:	depression, Beck Depression Inventory, normative data, confirmatory factor analysis, factorial structure

Beck Depression Inventory, Second Edition, Czech Version: Demographic Correlates, Factor Structure, and Comparison with Foreign Data

Marketa Ciharova,[a,e] Hynek Cígler,[b] Veronika Dostálová,[a] Gabriela Šivicová,[c] Ondrej Bezdicek,[a,d,f*]

[a] *Diamant Neuropsychology Laboratory, Department of Neurology and Centre of Clinical Neuroscience, 1st Faculty of Medicine and General University Hospital in Prague, Charles University, Prague, Czech Republic.*

[b] *Department of Psychology & Institute for Research on Children, Youth, and Family, Faculty of Social Studies, Masaryk University, Brno, Czech Republic.*

[c] *Department of Psychology, Faculty of Arts, Charles University, Prague, Czech Republic.*

[d] *Prague College of Psychosocial Studies, Prague, Czech Republic.*

[e] *Section Clinical Psychology, Faculty of Behavioural and Movement Sciences, Vrije Universiteit Amsterdam, Amsterdam, the Netherlands.*

[f] *National Institute of Mental Health, Klecany, Czech Republic.*

*****Ondřej Bezdíček, PhD., Department of Neurology and Centre of Clinical Neuroscience, Charles University in Prague, First Faculty of Medicine and General University Hospital in Prague, Kateřinská 30, 128 21 Praha 2, Czech Republic. +4202 2496 5524, +420 774 895 795 (phone), +4202 2492 2678 (fax), ondrej.bezdicek@gmail.com (e-mail)

ORCID ID http://orcid.org/0000-0002-5108-0181

Word Count: 3682

Beck Depression Inventory, Second Edition, Czech Version: Demographic Correlates, Factor Structure, and Comparison with Foreign Data

Abstract

Objectives: We aimed to examine whether demographic characteristics (i.e., sex, age, and education) correlate with total scores of the Czech version of the Beck Depression Inventory (BDI-II), understand the factorial structure of this scale, compare our results with findings of studies conducted in other countries and provide preliminary normative data for use in clinical practice.

Methods: Data of 450 participants were analyzed by using correlation analysis, non-parametric tests, and confirmatory factor analysis (CFA).

Results: Women, and participants with lower education, tended to score higher than men, and participants with higher education. There was no significant relationship between age and total scores. CFA confirmed two factors: cognitive-affective and somatic. Czech participants scored lower than participants in other studies. Preliminary normative data are presented in the form of percentile values for the whole sample and stratified according to gender and education level.

Conclusions: We recommend the usage of the BDI-II total score while taking into account also the cognitive-affective and somatic factor subscores. The comparison of our results with other foreign findings shows the need for the development of locally specific normative values for self-reported depression scales.

Keywords: depression, Beck Depression Inventory, confirmatory factor analysis, factorial structure, normative data

Introduction

The Beck Depression Inventory (BDI) is one of the most valid self-assessment methods for measuring depression severity (Beck, Ward, Mendelson, Mock, & Erbaugh, 1961; Lahlou-Laforêt, Ledru, Niarra, & Consoli, 2015; Richter, Werner, Heerlein, Kraus, & Sauer, 1998; Subica et al., 2014). Its second edition (BDI-II) consists of 21 items addressing cognitive, affective, motivational and physiological symptoms of depression, which may be used in adults and adolescents older than 13 years (Beck, Steer, & Brown, 1996). The BDI-II is uncomplicated to assign, cost-effective and highly convenient for clinical practice and research (Wang & Gorenstein, 2013). It evinces appropriate content validity in comparison to the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV; American Psychiatric Association, 2000), good sensitivity and moderate specificity, satisfactory test-retest reliability, and high internal consistency (Arbisi, 2001; Beck et al., 1996; Dozois, Dobson, & Ahnberg, 1998; Kjærgaard, Wang, Waterloo, & Jorde, 2014).

The BDI-II has been commonly used in many countries. As depressive symptoms may differ throughout various cultures and countries (American Psychiatric Association, 2000), and mental health measures may yield to culturally-relevant bias because of inappropriateness, different cultural meanings, and behaviour of participants (Van de Vijver & Tanzer, 2004), the development of local norms is necessary for evidence-based clinical practice (Wang & Gorenstein, 2013).

To provide valid results, normative values should reflect the associations between total scores and demographic variables. For instance, when differences are consistently found between men and women, sex-stratified norms may be needed for a given sample (Wang, & Gorenstein, 2013). Nevertheless, associations between total sum score and sex, age, and education were reported in previous research of the BDI-II often with contradicting results. For example, while some studies reported a significant difference between the BDI-II total score

for men and women, with women scoring higher than men (Al-Turkait & Ohaeri, 2010; Campos, & Gonçalves, 2011; Gomes-Oliveira, Gorenstein, Neto, Andrade, & Wang, 2012; González, Rodríguez, & Reyes-Lagunes, 2015; Kojima et al., 2002; Ptacek, Raboch, Vnukova, Hlinka, & Cervenkova, 2017; Roelofs et al., 2013; Sanz, Perdigón, & Vásquerz, 2003), other (Ghassemzadeh, Mojtabai, Karamghadiri, & Ebrahimkhani, 2005; Jakšić, Ivezić, Jokić-Begić, Surányi, & Stojanović-Špehar, 2013; Kapci, Uslu, Turkcapar, & Karaoglan, 2008) did not reveal any association between mean score and sex, or revealed the association only in samples of a certain nature (Aasen, 2001; Ginting, Näring, Van Der Veld, & Srisayekti, 2013). Age has been correlated inversely (Beck et al., 1996), positively (Jakšić et al., 2013; Sanz et al., 2003), or did not have any relationship with the total BDI-II score (Ghassemzadeh et al., 2005; González et al., 2015; Kojima et al., 2002; Roelofs et al., 2013). Moreover, the education level has also been found to correlate with the total score (Jakšić et al., 2013; Roelofs et al., 2013; Sanz et al., 2003).

Normative data for the Czech translation of the BDI-II in the Czech Republic have not yet been established, and, to the best of our knowledge, the two existing studies investigating psychometric properties of the BDI-II in Czech population (Ociskova et al., 2017; Ptacek et al., 2017), with the exception of sex, did not focus on demographics characteristics. Moreover, the only normative study of the BDI-II in a post-communist country was conducted in primary care settings in Croatia (Jakšić et al., 2013).

In the Czech Republic, estimations of lifetime prevalence of major depression vary between 7.8% (Andrade et al., 2003) and 13% (Dzúrová, Smolová, & Dragomirecká, 2000), while numbers for 12-month prevalence range from 2% (Andrade et al., 2003) to 4%. However, it has been argued that these, possibly underestimated, numbers do not account for patients institutionalized for their complaints as participants of these studies come from the general population (Formánek et al., 2019), and institutionalization is a

more common form of primary mental health care than in some other European countries where these estimations are lower (Winkler et al., 2017). Moreover, 12.5 out of 100 000 inhabitants of the Czech Republic terminate their lives by suicide (Czech Statistical Office, 2019). Efficient diagnostics and timely intervention are thus crucial, and quality instruments needed. Examination of demographic correlates of total scores of the BDI-II in the Czech Republic and development of normative data is hence of great importance for an evidence-based detection of depressive disorder.

In clinical practice, the comprehensive use of BDI-II scores is necessary (Abubakar et al., 2016). This is a reason for the analysis of the factorial structure of the BDI-II. Evidence may be found for models differing in the number of factors. Most research on the BDI-II factorial structure converges on two factors solution: cognitive-affective and somatic-vegetative, or somatic-affective and cognitive **(Huang and Chen, 2014; McElroy et al., 2018; Faro and Pereira, 2020)**. Both factors account for large amount of variance, such as 47.34% for the Croatian version or 42.61% for the Turkish version (Jakšić et al., 2013; Kapci et al., 2008). **A meta-analysis supported the two-factor structure, and it also suggested the existence of a general depression factor (and thus bifactor solution) as in other studies (Huang and Chen, 2014; McElroy et al., 2018; Faro and Pereira, 2020). Moreover, as more different models had similar fit to pooled data across several samples and language versions, there is not any conclusion about the exact factor structure. Unfortunately,** the authors of the Czech version of the BDI-II did not provide information on the factorial structure.

The primary goal of the present study was to analyze associations between the total BDI-II score and demographic variables. In addition, given the previous inconclusive results, we aimed to investigate the factorial structure of the Czech version of the BDI-II to understand the construct the BDI-II supposedly measures better. We also aimed to compare our results with findings from previous validation studies. Finally, regarding the lack of normative data for the

BDI-II Czech version, we wanted to provide preliminary normative standards for the BDI-II for its use in clinical practice.

Materials and Methods

Participants

The preliminary normative sample was collected using a cross-sectional design. **Participants (*N* = 616) were community residents recruited through convenience, non-random quota (defined as regions) sampling in all 14 regions of the Czech Republic in 2016, i.e., the sample was regionally representative, and via advertisements on the website of the General University Hospital in Prague and Prague School of Psychosocial Studies**. Regarding sample size estimation and the generalizability of the sample according to EFPA Review Model for the Description and Evaluation of Psychological and Education tests, see Appendix 1 (Evers et al., 2013). **Exclusion criteria were (a) age < 18 years, (b) uncorrected visual or hearing impairment, and (c) major depression or other major psychiatric disorder, a history of alcohol or substance abuse, epilepsy, previous head injury resulting in unconsciousness, any neurodegenerative disease, stroke, undergoing radiotherapy or chemotherapy and unstable medical illness (to ensure a healthy normative sample; Schmidt & Pardo, 2014). Due to fulfilling one or more exclusion criteria, 127 participants were not eligible.**

Additionally, 39 participants were excluded from the study **due to missing data on several items** in the BDI-II. Hence, 450 participants between the age of 18 to 96 ($M = 43.3$ years; **Median [*Md*] = 42; Interquartile range [*IQR*]** = 36) were included in the final sample (demographic characteristics are displayed in Table 1). For the confirmatory factor analysis (CFA), the data from 127 excluded participants were analyzed (see below) as well; their demographic characteristics are thus shown in Appendix 2.

[Table 1 near here]

Materials

Eligibility

A medical questionnaire, examining psychiatric diagnoses, neurological disorders, previous head injuries, and serious physical conditions, was administered to check the exclusion criteria.

Depressive symptoms

The BDI-II is a 21-item self-report questionnaire that measures the severity of depressive symptoms during the last two weeks, rated on a scale from 0 to 3, three being the most severe. The total score, ranging from 0 to 63 points, indicates depression severity (Beck et al., 1996). The BDI-II was administrated in the Czech language. The Czech version was published in 1999, its translation was four times revised and subsequently used in two pilot studies, one with a clinical sample and one with young soldiers with no mental disorder (Preiss & Vacíř, 1999). The BDI-II was administered according to the standard instructions used in the manual (Beck et al., 1996; Preiss & Vacíř, 1999).

Procedure

The participants signed an informed consent according to ethical standards and agreed on participation in the study. A copy of the informed consent with the description of the study was provided to the participant. First, the examinee was invited to fill in the medical status questionnaire. Then the BDI-II was administered.

Statistical Analyses

Statistical analyses were conducted using the Statistical Package for Social Sciences (IBM Corp., 2012). All tests were two-sided, with a significance level of .05. The evaluation of data distribution was conducted using Shapiro-Wilk tests ($p > .05$), analyses of skewness ($|z| < 1.96$) and kurtosis ($|z| < 1.96$), visual evaluation of histograms, Q-Q plots, and box plots. Considering the asymmetrical data distribution investigated using Shapiro-Wilk test ($p < .001$), different variance in sex, age and education groups according to Levene's test ($p < .001$) and ordinal data derived from a Likert-type scale, further statistical analyses were performed by non-parametric statistical methods. Spearman's correlation coefficient was used to examine the relationship between the BDI-II total score and age, between-groups differences were analyzed by the Mann-Whitney U test (men/women) or Kruskal-Wallis test (lower/higher/university education). Effect sizes were calculated using the equation suggested by Cohen (Fritz, Morris, & Richler, 2012) as $r = z/\sqrt{N}$, where z is the test statistic and N is the sample size.

The factorial structure of the BDI-II was assessed using ordinal CFA in the R environment with the lavaan package (Revelle, 2017). We fitted the one-dimensional and the two-factor model proposed by Storch, Roberti and Roth (2004), **however, there are alternative solutions, which we present in Appendix 3 (Huang and Chen, 2014). In the former, items 1–14, 17 and 21 constitute a cognitive-affective factor, while items 15, 16, 18–20 load on a somatic factor. This model was used also as a basis for bifactor solution, where all the items were loaded by general depression factor, while the definition of specific factors was the same as in Storch et al. (2017) model. We compared all models using difference fit indices as proposed by Cheung and Rensvold (2002) and using the robust likelihood-ratio test (Satorra and Bentler, 2010) if possible. All models were estimated with WLSMV estimator with robust correction, using the polychoric correlation matrix.**

To compare the mean of the Czech sample with the foreign data, the means of other normative and validation studies, as well as their confidence intervals, were investigated based on search in the literature. The confidence intervals, appearing as error bars in the figure, were counted using the number of participants in the sample and the standard deviation reported in the studies. Since our sample size did not meet sample size requirements for a normative study (Bridges & Holler, 2007), and the sampling was not random, we only provide preliminary normative data.

Results

Relationship between Demographic Variables and the BDI-II Total Score

In general, we found significant differences with small effect sizes between the BDI-II total score and sex (Mann-Whitney test $z = -1.97$, $p < .049$, $r = .01$) and education level (Kruskal-Wallis test $\chi^2(2) = 24.6$, $p < .001$), but no monotonic association of the BDI-II total score with age (Spearman's $rho = .084$, $p = .074$). The subsequent post-hoc analysis revealed that the differences were significant only for some education groups. Participants with lower education scored significantly higher than both groups of participants with higher education ($z = -3.96$, $p < .001$, $r = .19$) and a university degree ($z = -4.49$, $p < .001$, $r = .21$). No statistical difference was found between total scores of participants with higher education and a university degree ($z = -1.35$, $p = .177$, $r = .01$).

Confirmatory Factor Analysis and Internal Consistency

Table 2 shows all results for CFA; standardized factor loadings are displayed in Table 3. As the **Spearman's correlation matrices between both subsamples did not differ (Steiger test $\chi^2(210) = 218.17$, $p = .335$)** and model was invariant across groups (Appendix 4), we used the whole sample, including excluded participants (N = 577).

First, we investigated the factor structure of the BDI-II using the basic one-dimensional model by Storch, Roberti and Roth's (2004). This model fitted data poorly, but its internal consistency was good (Cronbach's $\alpha = .837$ and McDonald's $\omega_{hier} = .877$).

The two-dimensional Storch, Roberti and Roth's (2004) model fitted data significantly better ($p < .001$) and fit indices suggested its superiority according to the cuts proposed by Cheung and Rensvold (2002). Correlation of factors in our solution was high, $r = .772$ ($CI_{95\%} = [.718; .827]$) with high reliabilities: for the cognitive-affective factor $\alpha = .793$ and $\omega_{hier} = .850$, for somatic factor, $\alpha = .680$ and $\omega_{hier} = .678$ and $\omega_{tot} = .876$ for the whole questionnaire. **However, bifactor model fitted data even better, as shown in Table 3. On the other hand, reliabilities of specific factors were too low for practical assessment purposes.**

[Table 2 near here]

[Table 3 near here]

Comparison to Other Normative Studies

To investigate the mean, SD and range, descriptive statistical analyses were performed (Table 4). Descriptive statistics of the BDI-II investigated in the original study by Beck and his colleagues (1996) as well as validation studies adapting the instrument for different languages and culturally-specific groups are presented in Table 5. The studies are presented according to their geographical location. The results are depicted in Figure 1. When comparing the means and confidence intervals of the Czech and other studies, it may be observed that Czech participants scored significantly lower, except for the student sample in the Norwegian validation study (Aasen, 2001). The range of their answers was also narrower (0-32 vs. 0-62 or 0-51) than ranges reported in other studies.

[Table 4 near here]

[Table 5 near here]

[Figure 1 near here]

Preliminary Normative Data

Normative data in percentiles are presented stratified according to sex and education (Table 6), normative data for the whole sample can be found in Appendix 5. Since the analyses revealed no significant difference between the BDI-II total score of participants with higher education and a university degree, these two groups were merged in the development of normative data (i.e., lower/higher education).

[Insert Table 6 about here]

Discussion

In the present study, we investigated the influence of demographic characteristics (sex, age, and education) on the BDI-II total score. The results suggested that there was a relationship between the BDI-II total score and sex, as well as education, both results with small effect sizes. No association was found between age and BDI-II total score.

Women scored slightly higher in the BDI-II than men. This result is consistent with previous research, which reports higher scores endorsed by women than men (Aasen, 2001; Al-Turkait & Ohaeri, 2010; Beck et al., 1996; Campos, & Gonçalves, 2011; Gomes-Oliveira et al., 2012; González et al., 2015; Kojima et al., 2002; Ptacek et al., 2017; Roelofs et al., 2013; Sanz et al., 2003), and in accordance with the finding that depression and depressive symptoms are in general more often reported by women than men (American Psychiatric Association, 2013). On the other hand, some authors did not find any relationship between the scores of men and women. This may be explained by a different composition of their samples, which consisted of

students or patients, or possibly by an unbalanced number of participants regarding sex (Aasen, 2001; Ghassemzadeh et al., 2005; Jakšić et al., 2013; Kapci et al., 2008). In the present sample, the number of men and women was balanced. Another explanation is the combination of relatively small sample sizes and small effects in previous studies, which resulted in non-significant results due to sampling error.

Regarding the association of age with the BDI-II total score, the analyses did not reveal any significant relationship. These results are consistent with results of other normative and validation studies (Ghassemzadeh et al., 2005; González et al., 2015; Kojima et al., 2002; Roelofs et al., 2013). Although some authors reported a correlation between age and BDI-II total score, their samples again differed in their composition. Beck and colleagues (1996), who reported that age was inversely correlated with BDI-II total score, investigated this relationship on a student sample, whose age range was considerably narrower than ours (M = 19.58, SD = 1.84). While Jakšić et al. (2013) reported a positive correlation between age and the BDI-II total score in a medical outpatient sample (M = 55.01, SD = 12.99), our inclusion criteria were rather strict regarding medical conditions.

Regarding the influence of education, participants with lower education tended to score higher in the inventory and vice versa. These results are consistent with previous research (Jakšić et al., 2013; Roelofs et al., 2013; Sanz et al., 2003).

The result of our CFA revealed strong evidence speaking in favour of the BDI-II construct validity. This model is conceptually consistent with models by Al-Turkait and Ohaeri (2010) or de Miranda Azevedo et al. (2016), as well as Storch, Roberti, and Roth (2004), whose findings were confirmed by our analysis**. On the other hand, different model specifications are possible, as Huang and Chen (2014) showed, and therefore, we provided fit indices of these different models in supplemental materials. It is quite certain that BDI-II has a two-factorial structure (cognitive-affective vs. somatic) even in the Czech language, likely**

supplemented with a general depression factor. However, the exact specification of these two specific factors remains unclear. The suggested bifactor structure is in accordance with previous studies (McElroy et al., 2018; Huang and Chen, 2014; Faro and Pereira, 2020); item responses are based on the general factor (depression), while specific uncorrelated factors (the same as in two-factor solution) are present. The interpretations of two-factors and bifactor models are different. The two-factors model includes two closely related depression factors and the meaning of a total score is spurious. In the bifactor model, there is the general depression factor in addition to specific factors, while all the factors are orthogonal.

A comparison between the mean total score in the Czech study and other studies revealed that the scores of the Czech sample are significantly lower than all the other samples. **Our results are in line with previous studies, which reported lower 12-month prevalence of depression in the Czech Republic (2% - 4%; Andrade et al., 2003; Formánek et al., 2019) than was found in a meta-analysis of prevalence studies coming from multiple European countries (6.9%; Wittchen et al., 2011). Moreover, lower lifetime (7.8%), as well as 30-day (1%) prevalence of depression, was found in the Czech Republic than in multiple other countries (i.e., Brazil, Canada, Chile, Germany, Mexico, the Netherlands, the United States; ranges 8.1-16.9% and 1.3-4.6%, respectively; Andrade et al., 2003).**

This observation must be considered with caution, since our normative data of only a preliminary nature were, as mentioned above, collected in a non-random sampling design and our sample did not meet requirements for the size of a normative sample. Moreover, this may be again caused by the strict inclusion criteria of the current study as well as oversampling of young adults. Although there was no association between age and total sum scores in our sample, some authors (Jakšić et al., 2013; Sanz et al., 2003) reported that older participants tended to endorse higher scores in the BDI-II than younger participants. **This difference may**

also be influenced by cultural factors, since in different cultures depressive symptoms may be described in different terms (Van de Vijver, 2002), and thus, our results may reflect a problem with the translation (i.e., content validity problem) of the BDI-II that might lead to underreporting of symptoms. In addition, possible measurement error needs to be acknowledged, as samples may be very heterogeneous and measurements may vary across the studies. Undoubtedly, our results should be supported by further research, comprising bigger sample sizes and if possible also individual patient data meta-analyses, which could also examine how such heterogeneity may influence results. Overall, nevertheless, these findings corroborate the need for the development of pertinent normative values for each country rather than using norms developed for different countries.

The present study has several limitations. First, the normative study participants were chosen according to strict criteria to eliminate respondents with a psychiatric or neurological disorder as well as participants whose medical condition might increase the score in depression self-report measures. This may have resulted in one of the lowest average depression levels in the general population in comparison to other normative studies. Second, the sampling method was non-random, and it is questionable if such sampling can provide representative and normative data regarding the general population. Our criteria, however, seem to replicate the data collection and criteria used in other studies (Jakšić et al., 2013; Roelofs et al., 2013; Sanz et al., 2003). Third, since the concept of depression and personal burden of depressive symptoms differ throughout cultures and countries (Arbisi, 2001), and different historical milestones shaped other societies (Jakšić et al., 2013), in some of them, the BDI-II was adjusted regarding items formulation (Abubakar et al., 2016). It is not the case in the present study.

In conclusion, major depressive disorder and depressive symptoms as a comorbidity with other psychiatric diseases are in the Czech Republic a primary health concern in settings providing mental health services. Clearly, depressive symptoms have to

be assessed timely and effectively, and therefore, evidence-based, efficient and personalized diagnostic methods are needed. One of the most widely used diagnostic self-rating scales is the BDI-II. The present study replicates relationships between demographic variables and the BDI-II total score known from other studies. Furthermore, based on CFA, we argue for the diagnostic usage of BDI-II total score, while taking into account also the cognitive-affective and somatic factor subscores. In addition, the comparison of our results with other normative studies confirms the need for the development of locally or culturally specific normative values for self-reported depression scales. Finally, it provides preliminary BDI-II percentile values in the Czech version for use in clinical practice and meta-analytic efforts.

Key points

- Women scored higher in the BDI-II than men.
- Participants with lower education scored higher in the BDI-II than participants with higher education.
- CFA confirmed two factors: cognitive-affective and somatic.
- Preliminary normative data for the Czech version of the BDI-II are stratified according to gender and education.

Acknowledgements

The authors are grateful to Prague College of Psychosocial Studies for support regarding the place for the organization of the study and personal resources. The authors would like to thank for help with data collection to following students: Barboričová Sandra, Bystroňová Tereza, Čížková Jana, Galušková Karolína, Giunta Veronika, Havlík Filip, Hejlová Kristýna, Heřmánková Tereza, Holáreková Michaela, Holečková Hellen, Horák Jaroslav, Jančovičová Markéta, Juhaňák Jan, Kochová Kateřina, Kosová Jana, Kráčmarová Lucie, Kuldová Kateřina, Landa Vojtěch, Lojková Jiřina, Mýtina Ondřej, Plechač Vít, Popova Alina, Samková Nikola, Soósová Nina, Součková Markéta, Staníčková Anna, Šimůnková

Nina, Šteklová Kateřina, Švejnohová Iveta, Tulachová Karolína, Vašatová Veronika, Veverková Martina, Vintoniv Roksana, Vovsová Lucie, Wagner Daniel, Zborník Tadeáš Samuel, Zika Věra, Zoubek Martin.

Funding Details

This study was supported by the Czech Science Foundation (GA CR) ("Cognitive Predictors of Neurodegeneration," Grant Nr. 16-01781S and "Emotional Creativity and Cognitive Decline in the Elderly", Grant Nr. 18-26094S); by Prague College of Psychosocial Studies (Grant Nr. GA PVŠPS č. 18.0619.01); by Charles University, PROGRES Q27, by Ministry of Health of the Czech Republic, grant Nr. AZV 17-32445A; and by GA17-09797S ("Method studies for Czech versions of measures of psychological characteristics of adolescents and young adults").

Declaration of Interest Statement

The authors declare that they have no competing interests and that there are no disclosures to report.

The authors have no financial interests in the BDI-II Czech version.

Data Availability Statement

The dataset is available from the corresponding author on request.

References

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

Abubakar, A., Kalu, R. B., Katana, K., Kabunda, B., Hassan, A. S., Newton, C. R., et al. (2016). Adaptation and Latent Structure of the Swahili Version of Beck Depression Inventory-II in a Low Literacy Population in the Context of HIV. Plos One, 11, 1-15. doi: 10.1371/journal.pone.0151030

Al-Turkait, F. A., & Ohaeri, J. U. (2010). Dimensional and hierarchical models of depression using the Beck Depression Inventory-II in an Arab college student sample. BMC Psychiatry, 10, 1–14. doi: 10.1186/1471-244X-10-60

American Psychiatric Association (2000). Diagnostic and statistical manual of mental disorders: DSM-IV-TR. Washington, D.C.: American Psychiatric Association.

American Psychiatric Association (2013). Diagnostic and statistical manual of mental disorders (DSM-5®). Washington, D.C.: American Psychiatric Association.

Andrade, L., Caraveo-Anduaga, J. J., Berglund, P., Bijl, R. V., Graaf, R. D., Vollebergh, W., ... & Kawakami, N. (2003). The epidemiology of major depressive episodes: results from the International Consortium of Psychiatric Epidemiology (ICPE) Surveys. *International journal of methods in psychiatric research*, *12*(1), 3-21.

Arbisi, P. A. (2001). Beck Depression Inventory-II. In J.C. Impara, & B.S. Plake (Eds.). The mental measurements yearbook. Lincoln, Neb: Buros Institute.

Beck, A. T., Steer, R. A., & Brown, G. K. (1996). Beck depression inventory-II. San Antonio, 78, 490-498.

Beck, A. T., Ward, C. H., Mendelson, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. Archives of general psychiatry, 4, 561-571.

Bridges, A. J., & Holler, K. A. (2007). How many is enough? Determining optimal sample sizes for normative studies in pediatric neuropsychology. *Child Neuropsychology*, *13*(6), 528-538.

Campos, R. C., & Gonçalves, B. (2011). The Portuguese Version of the Beck Depression Inventory-II (BDI-II) Preliminary Psychometric Data with Two Nonclinical Samples. European journal of psychological assessment, 27, 258-264. doi: 10.1027/1015-5759/a000072

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. Structural equation modeling, 9, 233-255. doi: 10.1207/S15328007SEM0902_5

Czech Statistical Office (2019, December 10). *Sebevraždy* [Suicides]. Retrieved from https://www.czso.cz/csu/czso/sebevrazdy_zaj

de Miranda Azevedo, R., Roest, A. M., Carney, R. M., Denollet, J., Freedland, K. E., Grace, S. L., ... & de Jonge, P. (2016). A bifactor model of the Beck Depression Inventory and its association with medical prognosis after myocardial infarction. Health Psychology, 35, 614. doi: 10.1037/hea0000316

Dozois, D., Dobson, K., & Ahnberg, J. (1998). A psychometric evaluation of the Beck Depression Inventory–II. Psychological Assessment, 10, 83-89. doi: 10.1037/1040-3590.10.2.83

Dzúrová, D., Smolová, E., & Dragomirecká, E. (2000). *Duševní zdraví sociodemografických souvislostech* [Mental health in the sociodemographic context]. Prague: Charles University and the Prague Psychiatric Centre.

Evers, A., Hagemeister, C., Høstmælingen, A., Lindley, P., Muñiz, J., & Sjöberg, A. (2013). *EFPA review model for the description and evaluation of psychological and educational tests. Test review form and notes for reviewers.* Version 4.2, 6.

Faro, A., & Pereira, C.R. (2020). Factor structure and gender invariance of the Beck Depression Inventory – second edition (BDI-II) in a community-dwelling sample of adults. *Health Psychology and Behavioral Medicine, 8*(1), 16–31. doi:10.1080/21642850.2020.1715222

Formánek, T., Kagström, A., Cermakova, P., Csemy, L., Mlada, K., & Winkler, P. (2019). Prevalence of mental disorders and associated disability: Results from the cross-sectional CZEch mental health Study (CZEMS). *European Psychiatry*, *60*, 1-6.

Fritz, C. O., Morris, P. E., & Richler, J. J. (2012). Effect Size Estimates: Current Use, Calculations, and Interpretation. Journal of Experimental Psychology: General, 141, 2–18. doi: 10.1037/a0024338

Ghassemzadeh, H., Mojtabai, R., Karamghadiri, N., & Ebrahimkhani, N. (2005). Psychometric properties of a Persian-language version of the Beck Depression Inventory-second edition: BDI-II-Persian. Depression and Anxiety, 21, 185–192. doi: 10.1002/da.20070

Ginting, H., Näring, G., Van Der Veld, W. M., & Srisayekti, W. (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, 235–242. doi: 10.1016/S1697-2600(13)70028-0

Gomes-Oliveira, M. H., Gorenstein, C., Neto, F. L., Andrade, L. H., & Wang, Y. P. (2012). Validation of the Brazilian Portuguese Version of the Beck Depression Inventory-II in a community sample. Revista Brasileira de Psiquiatria, 34, 389–394. doi: 10.1016/j.rbp.2012.03.005

González, D. A., Rodríguez, A. R., & Reyes-Lagunes, I. (2015). Adaptation of the BDI-II in Mexico. Salud Mental, 38, 237–244. doi: 10.17711/SM.0185-3325.2015.033

Huang, C., & Chen, J.-H. (2014). Meta-Analysis of the Factor Structures of the Beck Depression Inventory–II. Assessment, 22(4), 459–472. doi: 10.1177/1073191114548873

IBM Corp. (2012). IBM SPSS Statistics for Windows (Version 21.0). [Software] Armonk, NY: IBM Corp.

Jakšić, N., Ivezić, E., Jokić-Begić, N., Surányi, Z., & Stojanović-Špehar, S. (2013). Factorial and Diagnostic Validity of the Beck Depression Inventory-II (BDI-II) in Croatian

Primary Health Care. Journal Of Clinical Psychology In Medical Settings, 20, 311-322. doi: 10.1007/s10880-013-9363-2

Kapci, E. G., Uslu, R., Turkcapar, H., & Karaoglan, A. (2008). Beck Depression Inventory II: evaluation of the psychometric properties and cut-off points in a Turkish adult population. Depression and Anxiety, 25, 104-110. doi: 10.1002/da.20371

Kjærgaard, M., Wang, C. A., Waterloo, K., & Jorde, R. (2014). A study of the psychometric properties of the Beck Depression Inventory- II, the Montgomery and Åsberg Depression Rating Scale, and the Hospital Anxiety and Depression Scale in a sample from a healthy population. Scandinavian Journal of Psychology, 55, 83-89. doi: 10.1111/sjop.12090

Kojima, M., Furukawa, T. A., Takahashi, H., Kawai, M., Nagaya, T., & Tokudome, S. (2002). Cross-cultural validation of the Beck Depression Inventory-II in Japan. Psychiatry Research, 110, 291-299. doi: 10.1016/S0165-1781(02)00106-3

Lahlou-Laforêt, K., Ledru, F., Niarra, R., & Consoli, S. M. (2015). Validity of Beck Depression Inventory for the assessment of depressive mood in chronic heart failure patients. Journal of affective disorders, 184, 256-260. doi: 10.1016/j.jad.2015.05.056

McElroy, E., Casey, P., Adamson, G., Filippopoulos, P., & Shevlin, M. (2018). A comprehensive analysis of the factor structure of the Beck Depression Inventory-II in a sample of outpatients with adjustment disorder and depressive episode. *Irish Journal of Psychological Medicine, 35*(1), 53-61. doi:10.1017/ipm.2017.52

Moore, T.M., & Haviland, M. (2010). Bifactor Models and Rotations: Exploring the Extent to which Multidimensional Data Yield Univocal Scale Scores. *Journal of Personality Assessment 92*(6), 544–559. doi:10.1080/00223891.2010.496477

Ociskova, M., Prasko, J., Kupka, M., Marackova, M., Latalova, K., Cinculova, A., …& Sigmundova, Z. (2017). Psychometric evaluation of the Czech Beck Depression

Inventory-II in a sample of depressed patients and healthy controls. *Neuroendocrinology Letters, 38*(2).

Preiss, M., & Vacíř, K. (1999). BDI-II. Beckova sebeposuzovací škála pro dospělé. [BDI-II. Beck's self-rating scale for adults.] Brno: Psychodiagnostika.

Ptacek, R., Raboch, J., Vnukova, M., Hlinka, J., & Cervenkova, M. (2017). Standardization of Czech version of Beck Depression Inventory (BDI-II). *European Psychiatry, 41*, S539.

Revelle W. (2017). Psych: Procedures for Personality and Psychological Research. Retrieved from https://cran.r-project.org/package=psych.

Richter, P., Werner, J., Heerlein, A., Kraus, A., & Sauer, H. (1998). On the validity of the Beck Depression Inventory. Psychopathology, 31, 160-168.

Roelofs, J., Van Breukelen, G., De Graaf, L., Beck, A., Arntz, A., & Huibers, M. (2013). Norms for the Beck Depression Inventory (BDI-II) in a large Dutch community sample. Journal of Psychopathology and Behavioral Assessment, 35, 93-98. doi: 10.1007/s10862-012-9309-2

Sanz, J., Perdigón, A. L., & Vásquerz, C. (2003). Adaptación española del Inventario para la Depresión de Beck-II (BDI-II): 2. Propriedades psicométricas em población general. [Spanish adaptation of the Beck's inventory for depression-II (BDI-II): 2. Psychometric properties in general population.] Clínica y salud: Revista de psicología clínica y salud, 14, 249-280.

Satorra, A., & Bentler, P. M. (2010). Ensuring positiveness of the scaled difference chi-square test statistic. Psychometrika, 75, 243-248. doi: 10.1007/s11336-009-9135-y

Schmidt, S., & Pardo, Y. (2014). Normative data. *Encyclopedia of Quality of Life and Well-Being Research*, *2014*, 4375-4379.

Storch, E. A., Roberti, J. W., & Roth, D. A. (2004). Factor structure, concurrent validity, and internal consistency of the beck depression inventory—second edition in a sample of college students. Depression and anxiety, 19, 187-189.

Subica, A. M., Fowler, J. C., Elhai, J. D., Frueh, B. C., Sharp, C., Kelly, E. L., & Allen, J. G. (2014). Factor structure and diagnostic validity of the Beck Depression Inventory–II with adult clinical inpatients: Comparison to a gold-standard diagnostic interview. Psychological Assessment, 26, 1106. doi: 10.1037/a0036998

Van de Vijver, F. J. R. (2002). Cross-Cultural Assessment: Value for Money? Applied Psychology, 51, 545-566.

Van de Vijver, F. J. R., & Tanzer, N. K. (2004). Bias and equivalence in cross-cultural assessment: an overview. Revue Européenne De Psychologie Appliquée/European Review of Applied Psychology, 54, 119-135. doi: 10.1016/j.erap.2003.12.004

Wang, Y. P., & Gorenstein, C. (2013). Psychometric properties of the Beck Depression Inventory-II: a comprehensive review. Revista Brasileira de Psiquiatria, 35, 416-431. doi: 10.1590/1516-4446-2012-1048

Winkler, P., Krupchanka, D., Roberts, T., Kondratova, L., Machů, V., Höschl, C., ... & Cerga-Pashoja, A. (2017). A blind spot on the global mental health map: a scoping review of 25 years' development of mental health care for people with severe mental illnesses in central and eastern Europe. *The Lancet Psychiatry*, *4*(8), 634-642.

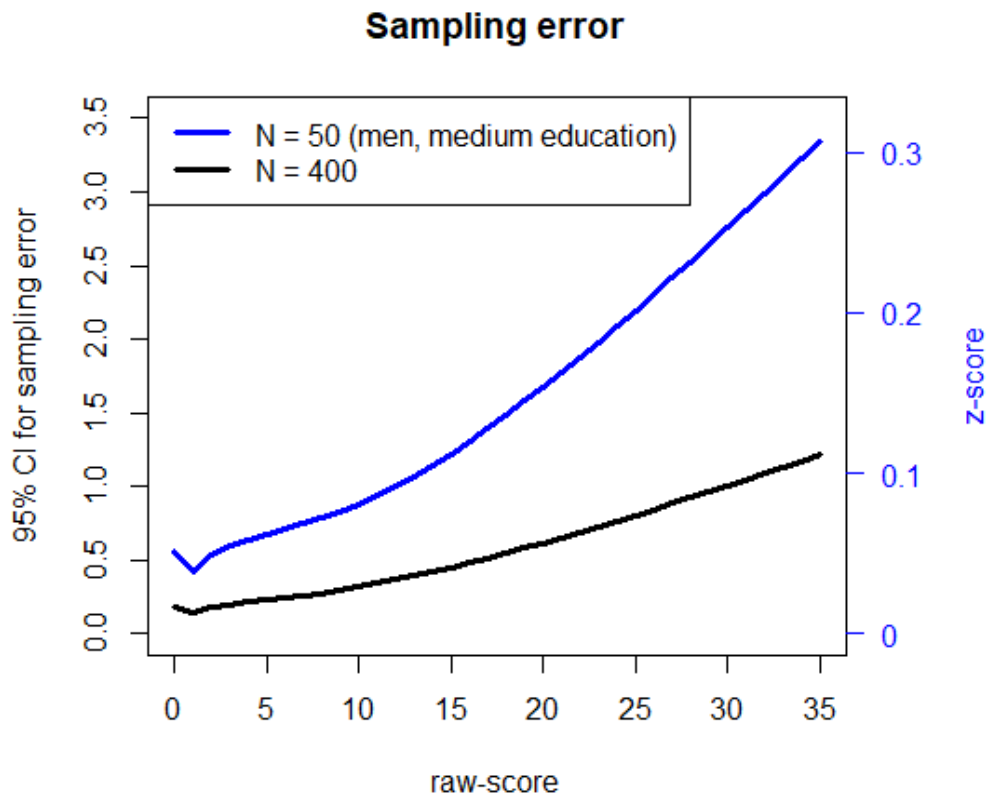
Wittchen, H. U., Jacobi, F., Rehm, J., Gustavsson, A., Svensson, M., Jönsson, B., ... & Fratiglioni, L. (2011). The size and burden of mental disorders and other disorders of the brain in Europe 2010. *European neuropsychopharmacology*, *21*(9), 655-679.

Wu, H., & Estabrook, R. (2016). Identification of confirmatory factor analysis models of different levels of invariance for ordered categorical outcomes. *Psychometrika, 81*(4), 1014–1045. doi:10.1007/s11336-016-9506-0

Appendices

Appendix 1

Sample size estimation and its generalizability



Note. Using the normative descriptives from a representative (regarding sampling and sample size with a total of $N = 7500$ respondents) study by Roelofs et al. (2012), and selecting from the whole sample a sample with $N = 400$ and subsample with $N = 50$ (we arbitrary chose descriptives for men with medium education level), the sampling error was still rather low as you can see on the attached plot. For example, the average raw score ($M = 10.6$) would have sampling error (its 95% confidence interval) about 1 score point (10% of standard deviation) if the sample size is only 50 respondents; if there was were about 400 respondents, the 95%

CI for the sampling error would be less than .33 raw-score points (3.3% of *SD*). On the other hand, for respondents with a very high score, the sampling error could be quite high, however, they are much above the usual cut-offs. We provide you with the analytical code to produce such a plot as our estimation approach could be unclear as Roelofs et al. (2012) used squared root of the raw score for norm estimation (and we back-transformed his estimates).

Appendix 2. Characteristics of Excluded Participants

Excluded participants (N = 127)		
Participants	*N*	%
Men	52	41
Women	75	59
18–30	22	17
30–40	9	7
40–50	25	20
50–60	25	20
60–70	18	14
70–96	28	22
Lower	4	3
Higher	73	58
University	50	39

Note. Lower education: Compulsory school attendance (9 years); Higher education: High school education; University: Obtained a degree.

Appendix 3. Alternative models

There are other alternative models describing BDI-II dimensionality. We used Huang and Chen (2014) models 2–4 and their bifactor counterparts 2a–4a from their meta-analysis. Models were specified exactly the same as in the original meta-analysis except for bifactor models, where we fixed latent covariances to zero; otherwise, the models are not identified (for comparison, see e.g. Moore and Haviland, 2010). Actually, Huang and Chen (2014) remarked that their models are not identified, but they did not provide any further explanation. We see that all the models fitted the data similarly well with the superiority of bifactor models.

Supplemental Table 2: Other possible factor solutions

	χ^2	df	p	TLI	RMSEA	90% CI		SRMR
model2	443,1	188	<.001	.931	.049	.043	.054	.081
model3	453,5	188	<.001	.928	.050	.044	.055	.083
model4	450.6	188	<.001	.929	.049	.043	.055	.083
model2a	258,9	168	<.001	.972	.031	.023	.038	.061
model3a	256,2	168	<.001	.973	.030	.023	.037	.061
model4a	260.6	168	<.001	.972	.031	.023	.038	.061

Note: See Huang and Chen (2014) meta-analysis for model specification (model names are the same).

Appendix 4. Measurement invariance across subsamples

We used the Storch's et al. (2004) two-dimensional model to compare both subsamples. The function measEq.syntax from semTools package was used for model construction with theta parameterization, and Wu and Estabrook's (2016) approach to identify residual variances. Configural model was estimated freely. In the metric model, both loadings and thresholds were constrained. Then item intercepts, item residual variances, latent means, latent variances, and finally, latent covariances were gradually constrained across subsamples.

As the fully constrained model fitted data the best and none of the likelihood-ratio tests were significant, we merged both subsamples. Then all the following analyses were performed on the whole sample.

							90% CI			LRT		
	x2	df	p	CFI	TLI	RMSEA	lower	upper	SRMR	$\Delta\chi^2$	Δdf	p
configural	625.1	376	<.001	0.930	0.922	0.048	0.041	0.055	0.103			
metric	634.5	395	<.001	0.933	0.928	0.046	0.039	0.052	0.104	20.6	19	.358
scalar	632.8	412	<.001	0.938	0.937	0.043	0.036	0.050	0.104	14.8	17	.609
residual	655.7	433	<.001	0.937	0.939	0.042	0.036	0.049	0.107	29.8	21	.096
means	670.4	435	<.001	0.934	0.936	0.043	0.037	0.050	0.106	5.3	2	.070
variances	659.8	437	<.001	0.937	0.94	0.042	0.035	0.049	0.107	2.1	2	.351
covariances	644.7	438	<.001	0.942	0.944	0.041	0.034	0.047	0.107	0.4	1	.529

Appendix 5. Normative data (percentiles) for the BDI-II (whole normative sample, N = 450)

BDI-II Total	Frequency	Cumulative Relative Frequency	Percentile
0	45	0.1	5
1	30	0.17	13
2	36	0.25	21
3	39	0.33	29
4	38	0.42	38
5	31	0.49	45
6	17	0.52	51
7	34	0.60	56
8	25	0.66	63
9	21	0.70	68
10	19	0.74	72
11	21	0.79	77
12	23	0.84	82
13	15	0.88	86
14	10	0.90	89
15	6	0.91	91
16	8	0.93	92
17	8	0.95	94
18	5	0.96	95
19	5	0.97	96
20	1	0.97	97
21	2	0.98	97
22	1	0.98	98
23	3	0.98	98
24	3	0.99	99
25	0	0.99	99
26	1	0.99	99
27	0	0.99	> 99
28	0	0.99	> 99
29	1	1.00	> 99
30	0	1.00	> 99
31	1	1.00	> 99
≥ 32	1	1.00	> 99

Note. BDI-II: Beck Depression Inventory, Second Edition. Percentile values were rounded to an integer.

Table 1. Participant characteristics

Normative Sample (N = 450)		
	Participants Number	(%)
Sex		
Men	206	46
Women	244	54
Age		
18–30	167	37
30–40	41	9
40–50	54	12
50–60	79	18
60–70	58	13
70–96	51	11
Education		
Lower	55	12
Higher	225	50
University	170	38
Race		
Caucasian	449	99.77
Asian	1	.23

Note. Lower education: Compulsory school attendance (9 years); Higher education: High school education; University: Obtained a degree.

Table 2. Comparison of confirmatory factor analyses (all participants, N = 577)

		fit indices					comparison to the 1-factor model			
model	name	df	χ^2	TLI	RMSEA [$CI_{90\%}$]	SRMR	Δdf	$\Delta\chi^2$	ΔTLI	ΔSRMR
1	1-factor	189	578.7	0.895	.060 [.054; .065]	0.094				
2a	2-factor (Storch et al.)	188	486.6	0.920	.052 [.047; .058]	0.088	1	64.5	.025	.006
3	bifactor (Storch et al.)	168	258.3	0.973	.031 [.023; .038]	0.061	21	158.2	.078	.033

Note. All the tests were significant on $p < .001$.

Table 3. Factor loadings of the confirmatory factor models (N = 577)

	1-factor	2-factor		bifactor		
	BDI	C-A	SOM	BDI	C-A	SOM
I1	.707***	.738***		.642***	.352***	
I2	.611***	.641***		.490***	.455***	
I3	.530***	.562***		.293***	.716***	
I4	.679***	.703***		.716***	.068	
I5	.513***	.542***		.342***	.555***	
I6	.507***	.535***		.335***	.551***	
I7	.686***	.717***		.544***	.526***	
I8	.576***	.603***		.483***	.380***	
I9	.444***	.477***		.280**	.540***	
I10	.544***	.564***		.491***	.274***	
I11	.657***	.683***		.651***	.186**	
I12	.541***	.559***		.537***	.144*	
I13	.626***	.652***		.597***	.232***	
I14	.663***	.688***		.605***	.315***	
I15	.686***		.753***	.623***		.591***
I16	.533***		.585***	.503***		.338***
I17	.647***		.699***	.727***		-.084
I18	.418***		.453***	.439***		.062
I19	.675***		.731***	.702***		.137*
I20	.710***		.777***	.654***		.460***

I21	.446***		.500***	.409***		.404***
ω_h		.831	.744	.690	.286	.126
ω_t	.877		.882		.884	

Note. BDI: Beck Depression Inventory; C-A: Cognitive-Affective Factor; I1 – I21: Items 1 through 21; S: Somatic Factor. In the 2-factor model, correlation of factors C-A and SOM was $r = .772^{***}$.

$^{*}p < .05$, $^{**}p < .01$, and $^{***}p < .001$.

Table 4. Descriptive statistic and psychometric properties of the normative sample, N = 450.

Mean	7.04
Standard deviation	5.81
Variance	33.711
Range	0 - 32

Note. Mean – 95% CI [6.51, 7.58]; Standard deviation - 95% CI [5.41, 6.17]; Variance - 95% CI [29.30, 38.12].

Table 5. Summary of validation studies for different countries and languages (Mean, Standard Deviation).

Author, year	Language	Country	N	Sample	M	SD	Range
Beck, Steer and Brown (1996)	English	USA	120	Student	12.56	9.93	NR
Roelofs et al. (2013)	Dutch	Netherlands	7500	Adult	10.6	10.9	0 – 62
Jakšić et al. (2013)	Croatian	Croatia	314	Adult	10.35	10.27	0 – 51
Arnarson et al. (2009)	Icelandic	Iceland	1206	Student	8.80	7.82	NR
Aasen (2001)	Norwegian	Norway	875 303	Adult Student	8.12 7.12	7.5 6.0	NR
Sanz et al. (2003)	Spanish	Spain	470	Adult	9.4	7.7	NR
Campos & Gonçalves (2011)	Portuguese	Portugal	538	Student	8.88	7.85	NR
Kapci et al. (2008)	Turkish	Turkey	153 50	Women Men	14.99 14.09	9.19 9.72	NR
González et al. (2015)	Spanish	Mexico	420 220	Student Adult	9.31 9.82	7.84 7.70	NR
Gomes-Oliveira et al. (2012)	Portuguese	Brazil	182	Adult	9.87	10.71	NR
Ghassemzadeh et al. (2005)	Persian	Iran	125	Student	9.79	7.96	NR
Al-Turkait & Ohaeri (2010)	Arabic	Kuwait	624	Student	15.5	8.5	NR
Kojima et al. (2002)	Japanese	Japan	766	Adult	8.90	6.52	NR
Ginting et al. (2013)	Bahasa Indonesia	Indonesia	720	Adult	14.20	9.7	NR
Abubakar et al. (2016)	Swahili	Kenya	221	Adult	18.20	8.06	NR

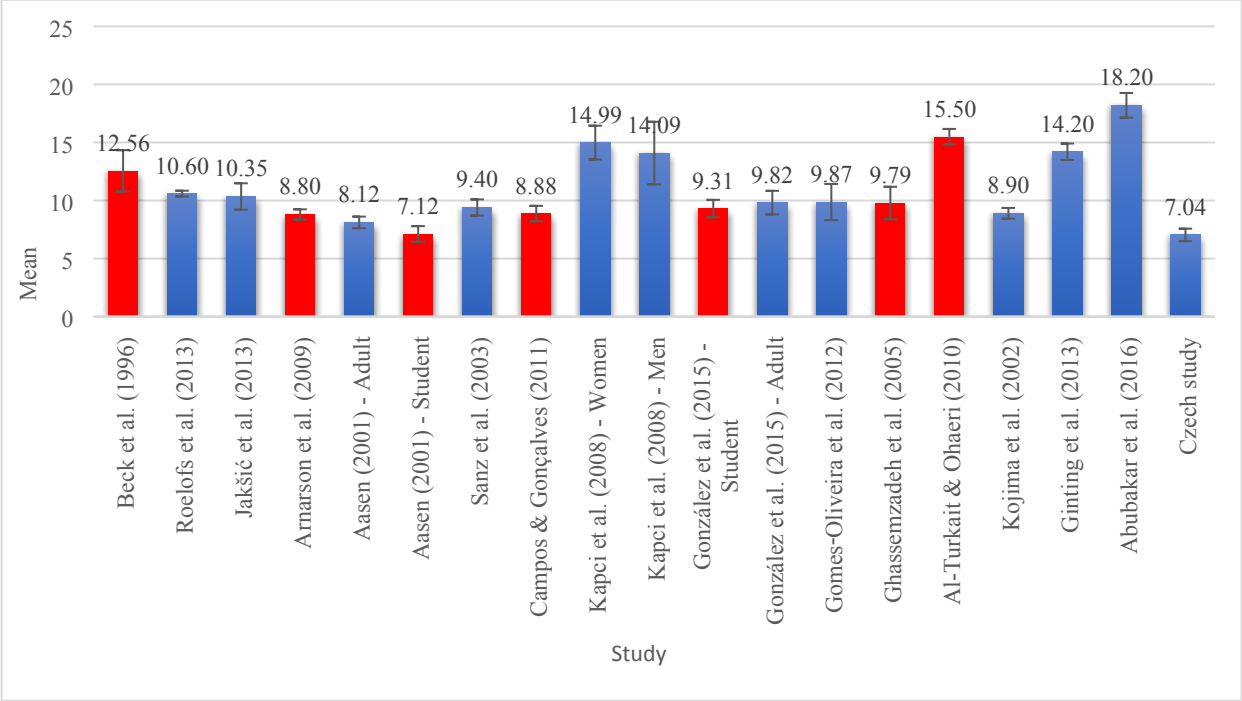
Note. M: Mean; NR: Not reported; SD: Standard Deviation.

Table 6. Normative standards (percentiles) for the BDI-II according to gender and education (whole normative sample, N = 450)

Gender	Women		Men	
Education	lower	higher	lower	higher
Mean (SD)	13.42 (8.541)	6.86 (5.204)	8.97 (6.531)	6.01 (5.323)
Median	11.50	6	8	5
Percentiles				
> 99	≥ 32	≥ 26	≥ 31	≥ 24
90–99	27–31	14–25	16–30	13–23
80–89	23–26	12–13	14–15	10–12
70–79	18–22	10–11	13	9
60–69	14–17	8–9	11–12	7–8
50–59	12–13	6–7	8–10	5–6
40–49	11	5	7	4
30–39	8–10	4	4–6	3
20–29	6–7	3	3	2
10–19	3–5	1–2	2	1
< 10	0–2	0	0–1	0

Note: SD: Standard Deviation. Lower education: Compulsory school attendance (9 years); Higher education: High school education; University: Obtained a degree. Percentile values were rounded to an integer.

Figure 1. Comparison of means in different studies.



Note. Red colour = Student samples. Blue colour = Adult samples.

Figure captions

Figure 1. Comparison of means in different studies.