Taylor & Francis
Taylor & Francis Group

ORIGINAL ARTICLE

Psychometric characteristics of the Beck Depression Inventory-II with college students of diverse ethnicity

DENNIS P. CARMODY

Department of Pediatrics, Robert Wood Johnson Medical School, University of Medicine and Dentistry of New Jersey, New Brunswick, NJ, USA

Abstract

Introduction: To evaluate the psychometric characteristics of the Beck Depression Inventory – Second Edition (BDI-II) with an ethnically diverse sample of college students. *Methods:* The BDI-II was completed by a group of 502 college students (54% women) with an ethnic distribution of African-American ($n=49$, 10%), Asian-American ($n=33$, 7%), Hispanic ($n=113$, 22%), Native-American ($n=10$, 2%), and White ($n=297$, 59%). Psychometric characteristics of the inventory with the ethnically diverse group were compared to the results published in the test manual for nonclinical samples composed predominantly of European and White participants. *Results:* Using confirmatory factor analyses, a three-factor model that identified negative attitude, performance difficulty, and somatic dimensions, provides a better fit of the data than does the two-factor model. Similar psychometric characteristics were found between the ethnically diverse student sample and the standardization sample. Based on multivariate analysis of variance, White students had higher scores on the item of agitation compared to Asian-American students and on the items of worthlessness and irritability compared to Hispanic students. *Conclusion:* The results of the psychometric analyses suggest that the BDI-II is suitable as a screening instrument for depression in college populations of diverse ethnicity.

Key Words: *Depression, screening instrument, psychometrics, confirmatory factor analysis, major depression inventory*

Introduction

Major depression is underdiagnosed and under-treated in the general populations of several countries in Asia, Europe, and North America [1]. Eight percent of adults will experience major depression and an additional 1% will experience bipolar disorder [2]. The general medical sector is the only source of care for one-half of all depressed patients; however, primary care physicians do not recognize depression in about one-half of the affected patients [3,4], and only 10% of patients are referred from general practice to psychiatrists for further evaluation [5]. Even when general medical providers recognize the presence of depression, many are hesitant to report it in the record due to concern about confidentiality, stigmatization, or reimbursement rates for the care of psychiatric as compared to physical conditions [6]. In contrast, depression may be over diagnosed in mental health settings where reimbursement is tied to mental health conditions [7]. In general, psychiatric disorders are under-estimated as contributors to disability with the costs of loss of years of productive life [8].

It has been shown that the prevalence of depression is related to gender, social class and culture. Major research studies have found that women have twice the risk as men for depression in developed countries, while the ratio varies in developing countries [9]. This 2:1 ratio has been confirmed in a large-scale study by the World Health Organization gathering data at 15 centers in four continents, suggesting that biological and social factors have similar effects across cultures [10]. Although the gender difference is age specific, with small differences in childhood or old age and notable differences in middle age [11], the incidence of severe affective disorder increases with age in both women and men [12]. In contrast, the relationship between ethnicity and depression is complex and influenced by many factors [13], including socioeconomic status [14,15]. In addition, given the cultural variations in the clinical presentations of depression [16,17] there is a need for culturally sensitive assessment

Correspondence: Dennis P. Carmody, PhD, Institute for the Study of Child Development, Robert Wood Johnson Medical School, University of Medicine and Dentistry of New Jersey, 97 Paterson Street, New Brunswick, NJ 08903, USA. Tel: +1-732-235-7700. Fax: +1-732-235-6189. E-mail: carmoddp@umdnj.edu

methods [18] that avoid the problem of category fallacy [19].

Assessments of depression include self-report inventories that are easily administered and rapidly scored. Among the popular self-report inventories of depression, the Beck Depression Inventory (BDI) [20] and the Beck Depression Inventory – Second Edition (BDI-II) [21] are widely used to measure depression in adults and adolescents [22,23]. Based on research with the original BDI, Beck indicated that depression is composed of an affective-cognitive factor and a somatic factor [20]. Beck and colleagues continued to use the same two-factor model to describe the 21 items of the BDI-II [21], with 16 items loading on the cognitive-affective factor and five items loading on the somatic factor.

The two-factor model was found in Canadian sample [24], a college sample [25], and a community sample completing the Japanese language version of the BDI-II [26], although there were differences among the studies in which 21 of the items were associated with the affective-cognitive and the somatic factors. Alternatively, there is support for a three-factor model, with the factors defined as negative attitude, performance difficulty, and somatic elements [27–30].

The participants in the psychometric studies of the BDI-II were primarily White, European American or North American. The college students in the Beck standardization sample were described as mostly White [21], and only 8% of the adult outpatients in a study of the construct validity of the BDI-II were minorities [31]. Given the current ethnic distributions of urban areas, there is a clear need to assess the psychometric properties of the English language version of the BDI-II need to be assessed with a multiethnic sample, one that represents the ethnic diversity of urban areas.

The purpose of this study was to determine how well the psychometric properties of the BDI-II compare to the published standards when used with a multiethnic, urban population. Specifically, do gender differences found in standardization samples with White populations exist in multiethnic populations? Do the factor compositions of items on BDI-II based on White student samples agree with item loadings based on a multiethnic, urban population? Does the underlying structure of the data from the multi-ethnic sample fit a two-factor or three-factor model of depression? To answer the research questions, a confirmatory factor analysis was conducted to assess the goodness of fit of these data to models of depression reported in the literature. The data from the study with a multiethnic population were fitted to the pattern matrices of a two-factor model [21] and a three-factor model [28] and the patterns were compared.

Methods

Participants

A total of 532 students at a private, urban college participated by volunteering through the subject pool in return for credit in their introductory psychology courses. Thirteen had incomplete inventories and 17 others did not indicate gender or ethnicity, leaving 502 participants (272 women, 230 men, 52.4% women, age range 17–46 years, mean 20.18 years, mode 20 years) with complete data sets. Ethnic identity based on self-report was distributed as African-American ($n = 49$, 10%), Asian-American ($n = 33$, 7%), Hispanic ($n = 113$, 22%), Native-American ($n = 10$, 2%), and White ($n = 297$, 59%) [The distribution of ethnicity in the college sample was compared to the ethnic distribution in the geographical region within 25 miles of the college, as described in USA Census 2000. Based on the race characteristics of several counties in northern New Jersey, the ethnic distribution is African-American (17%), Asian-American (11%), Hispanic (30%), Multi-racial (3%), Native-American (<1%), and White, not Hispanic (38%). Therefore, the college sample was a fair representation of the population in Northern New Jersey.]

Treatment of participants was in accordance with the ethical standards of the American Psychological Association, the study was performed in accordance with the ethical standards indicated in the Declaration of Helsinki [32], and the local Institutional Review Board approved the project. Each volunteer had the right to refuse participation, terminate the testing session, or not complete the inventory. The experimenter read the purpose and instructions for the experiment to participants, reviewed the consent form, and answered questions. After signing the consent form, the participants completed a packet containing the demographic form and the BDI-II.

Instrument

The BDI-II is a 21-item scale that measures self-reported depression in adults and adolescents aged 13 years and older [21]. Participants rated themselves on the 21 items of depressive symptoms by endorsing the most characteristic statements covering a time frame of the past 2 weeks. Statements range in intensity of symptoms in four steps from minimal to severe.

Data analyses

The statements chosen by participants determined points for each item. Items were scored as zero points for the mildest level of the symptom and as 1, 2, or 3 points for increasing levels of symptom severity. By adding the points for the 21 items, the total scores ranged from zero to 63 points. Using

Cronbach's α, analyses of the internal consistency of the BDI-II with the college sample were performed and the results were compared to the psychometric characteristics of student samples reported in the BDI-II manual [21]. Gender differences were examined by a comparison of total scores for women and men, and the results were compared to the Beck sample. A multivariate analysis of variance was applied to the responses to the 21 items to determine if there were differences in item endorsement due to gender and ethnicity.

A confirmatory factor analysis [33] was applied to the data to determine the fit of the data to the factor structures reported in the literature. Several recommendations in the design and analyses of a confirmatory factor analysis (CFA) study were followed [34]. The sample size exceeds 100 subjects, complete data are available for all subjects, variables are normally distributed, and the correlation or covariance matrix is reported.

Several models were tested on the multiethnic sample including the null-model, the one-factor model, the two-factor model, and the three-factor model. *Null model*: This model assumes that the 21 BDI-II items represent independent factors. The results are used as a baseline to compare the findings from the other models. *One-Factor Model*: This model indicates that the 21 items on the BDI-II load on a single factor that accounts for all symptoms of depression. The factor variance was set to 1.0. *Beck Two-Factor Model*: This model is based on the exploratory factor analysis described in the test manual [10]. Sixteen items were constrained to the cognitive-affective factor and five items were constrained to the somatic factor. The variance of each factor was set at 1.0 and the factors were correlated in the analysis. *Three-Factor model*: This model is based on the three-factor model reported by Osman [28]. The three factors were correlated in the analysis. Ten items were constrained to the negative attitude factor (BDI-II items 1, 2, 3, 5, 6, 7, 8, 9, 10, 14), seven items to the performance difficulty factor (BDI-II items 4, 12, 13, 15, 17, 19, 20), and four items to the somatic elements factor (BDI-II items 11, 16, 18, 21). The variance of each factor was constrained to 1.0 and three error covariances were included (item 6 with item 8, item 15 with item 20, and item 16 with item 20) to simulate the model by Osman.

The goodness of fit of each model was tested using the statistical criteria of the chi-square goodness-of-fit, the root mean square error of approximation (RMSEA), and the Comparative Fit Index (CFI) [33,34]. Using the statistical measures, the models were compared to determine which model has the best fit to the data. Better fitting models have lower chi-square values, lower RMSEA levels, and higher values of CFI.

Results

Descriptive statistics

The means, standard deviations, and item-total correlations are shown in Table I. The Cronbach coefficient α of 0.916 for the BDI-II items showed high internal consistency. Item responses showed item-total correlations in excess of 0.46 for 95% of

Table I. Descriptive statistics for the items on the BDI-II for the multiethnic college sample.

Item	Mean	Std. Dev	Skewness	Kurtosis	R
1	0.54	0.72	1.47	2.38	0.68
2	0.54	0.69	1.27	1.59	0.60
3	0.51	0.72	1.26	0.95	0.56
4	0.59	0.68	0.95	0.66	0.55
5	0.63	0.65	0.86	1.00	0.53
6	0.48	0.86	1.86	2.53	0.61
7	0.55	0.85	1.37	0.82	0.64
8	0.77	0.83	0.92	0.29	0.67
9	0.28	0.54	2.28	6.65	0.55
10	0.54	0.92	1.63	1.48	0.58
11	0.65	0.72	1.01	0.91	0.59
12	0.55	0.70	1.23	1.48	0.61
13	0.66	0.79	1.22	1.23	0.55
14	0.28	0.65	2.43	5.42	0.70
15	0.77	0.67	0.62	.057	0.59
16	1.04	0.79	0.35	−0.39	0.47
17	0.63	0.73	0.97	0.53	0.64
18	0.84	0.81	0.87	0.46	0.52
19	0.80	0.77	0.67	−0.09	0.69
20	0.82	0.74	0.71	0.36	0.62
21	0.26	0.60	2.67	7.52	0.36
Total Score	12.75	9.07			

$N = 502$. Items 9, 14, 21 have skewness problems and kurtosis problems.
R, item-total correlation.

the items and greater than 0.53 for 90% of items; the low correlation of 0.36 was found for the item of "Loss of Interest in Sex". As shown in Table I, 18 of the 21 BDI-II items had skewness values less than 2.0 and kurtosis values less than 3.0. The three items that had large values were item 9, suicidal thoughts (skewness 2.28, kurtosis 6.65), item 14, worthlessness (skewness 2.43, kurtosis 5.42), and item 21, loss of interest in sex (skewness 2.67, kurtosis 7.52). The intercorrelations among the items on the BDI-II are shown in Table II.

Comparisons were made between the multiethnic college sample and the Beck [21] student sample. The college sample ($M = 20.14$, SD $= 2.831$) was significantly older than the Beck sample ($M = 19.58$, SD $= 1.84$) by 0.56 years ($t_{498} = 4.452$; $P < 0.001$; 95% CI, 0.32–0.81). There was no difference in gender distribution between the college sample (54.2%) and the Beck sample (56% women) ($t_{501} = 0.816$; ns; 95% CI, -0.03 to 0.06), and the mean depression score for the college sample ($M = 12.75$, SD $= 9.07$) was similar to the Beck sample ($M = 12.56$, SD $= 9.93$) ($t_{501} = 0.462$; ns; 95% CI, -0.608 to $+0.982$).

Gender differences were compared. Women had higher depression scores ($M = 13.56$ SD $= 8.87$) than men ($M = 11.79$, SD $= 9.23$), ($t_{500} = 2.19$, $P < 0.05$). Depression scores for women in the college sample were similar to depression scores for women in the Beck sample ($M = 14.55$, SD $= 10.74$) ($t_{271} = 1.843$; ns; 95% CI, -2.050 to 0.068), while men in the college sample had higher depression scores than men in the Beck sample ($M = 10.04$, SD $= 8.23$) ($t_{229} = 2.869$; $P < 0.004$; 95% CI, 0.547–2.947).

Multivariate analyses of items

A 2×5 multivariate analysis of variance (MANOVA) was conducted to determine whether the scores on the 21 items differed significantly as a function of gender or ethnicity. Using Wilk's criteria, there was a significant main effect for gender ($F(21,472) = 1.684$, $P < 0.05$). Post-hoc univariate F-tests revealed that women had higher scores than men for item 1, sadness ($F(1,492) = 5.96$, $P < 0.02$), item 10, crying ($F(1,492) = 5.74$, $P < 0.02$), item 15, energy ($F(1,492) = 4.24$, $P < 0.05$), and item 20, fatigue ($F(1,492) = 5.43$, $P < 0.05$).

There was a significant main effect of ethnicity ($F(84,1900) = 1.469$, $P < 0.005$), with contributions from item 11, agitation ($F(4,492) = 2.95$, $P < 0.02$), and item 17, irritability ($F(4,492) = 2.47$, $P < 0.05$). Post-hoc Tukey HSD tests revealed that White students had higher scores on three items, specifically, higher agitation scores than Asian-American students ($P < 0.005$), higher worthlessness scores than Hispanic students ($P < 0.05$), and had higher irritability scores than Hispanic

students ($P < 0.01$). The interaction of gender and ethnicity was not reliable ($F(84,1900) < 1.0$, $P > 0.05$).

Confirmatory factor analyses

The fit indices for the models are presented in Table III. Both the two-factor and three-factor models met the criteria of CFI greater than 0.85 and RMSEA equal to or less than 0.10. The one-factor model did not meet the RMSEA criterion. By examining the differences in the chi-square values between the models, it was found that the two-factor model was a better fit than the one-factor model ($\triangle \chi^2 = 107.69$; df $= 1$; $P < 0.01$). The three-factor model was superior in fit to the two-factor ($\triangle \chi^2 = 70.07$; df $= 5$; $P < 0.01$). In summary, using the fit criteria of CFI and RMSEA, the three-factor model was the best fit to the data set.

Discussion

The psychometric characteristics of the BDI-II with an ethnically diverse population show high reliability. Women in the multiethnic college sample had higher scores than men, a finding that replicated gender differences reported in the standardization sample [21]. A confirmatory factor analyses was used to examine the sample data to determine the best fitting model. The two-factor model fit the data better than a one-factor model, providing support for the two dimensions of depression of cognitive-affective and somatic elements [21]. The three-factor model was a better fit of the data than the two-factor solution.

There have been reports of differential endorsements of items on the Beck Depression Inventory by minority groups, especially for Hispanic respondents [35]. However, in this study, no differences were found among ethnic groups for overall scores. The gender differences in levels of depressive symptoms typically reported in assessments of the BDI were preserved for all ethnic groups. The only differences found between ethnic groups were the differential item endorsements for three items. White respondents had higher scores on the items of agitation, worthlessness, and irritability. The similarity in depression scores across ethnic groups requires explanation. One possibility is that the sample size was too small and the tests chosen lacked statistical power to detect subtle differences. However, the sample sizes for ethnic groups were large, and the power of the statistical tests was sufficient to detect differences. Therefore, an explanation for the similarity in depression scores lies in the common culture of daily urban life experienced by the student sample. Although the concept of category fallacy needs to be examined when conducting cross-cultural research [19], the students in this sample,

Table II. Intercorrelations among the items of the BDI-II in the multiethnic college sample.

Item	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
1. Sadness	1.000																				
2. Pessimism	0.369	1.000																			
3. Past failure	0.357	0.409	1.000																		
4. Loss of pleasure	0.329	0.314	0.257	1.000																	
5. Guilty feelings	0.361	0.288	0.262	0.253	1.000																
6. Punishment feelings	0.400	0.352	0.348	0.298	0.283	1.000															
7. Self-dislike	0.413	0.423	0.469	0.362	0.297	0.354	1.000														
8. Self-criticalness	0.419	0.370	0.429	0.286	0.350	0.392	0.396	1.000													
9. Suicidal thoughts or wishes	0.422	0.301	0.315	0.319	0.298	0.308	0.342	0.350	1.000												
10. Crying	0.457	0.219	0.182	0.272	0.344	0.367	0.296	0.367	0.310	1.000											
11. Agitation	0.387	0.340	0.269	0.271	0.338	0.309	0.316	0.410	0.252	0.328	1.000										
12. Loss of interest	0.361	0.331	0.281	0.402	0.264	0.334	0.327	0.330	0.408	0.275	0.301	1.000									
13. Indecisiveness	0.288	0.322	0.243	0.210	0.242	0.275	0.331	0.374	0.240	0.295	0.255	0.308	1.000								
14. Worthlessness	0.437	0.483	0.538	0.307	0.276	0.461	0.488	0.503	0.407	0.285	0.362	0.424	0.316	1.000							
15. Loss of energy	0.396	0.254	0.234	0.273	0.257	0.289	0.305	0.320	0.330	0.263	0.248	0.380	0.280	0.376	1.000						
16. Changes in sleeping patterns	0.251	0.164	0.121	0.175	0.227	0.185	0.229	0.270	0.202	0.266	0.334	0.206	0.262	0.197	0.292	1.000					
17. Irritability	0.441	0.372	0.265	0.258	0.292	0.279	0.330	0.425	0.291	0.334	0.455	0.429	0.336	0.386	0.393	0.265	1.000				
18. Changes in appetite	0.305	0.219	0.201	0.251	0.189	0.295	0.259	0.311	0.149	0.208	0.302	0.240	0.283	0.319	0.296	0.320	0.290	1.000			
19. Concentration difficulty	0.377	0.390	0.367	0.369	0.307	0.330	0.372	0.470	0.320	0.377	0.377	0.405	0.441	0.474	0.426	0.299	0.398	0.317	1.000		
20. Tiredness or fatigue	0.389	0.298	0.238	0.319	0.267	0.295	0.311	0.317	0.223	0.276	0.319	0.336	0.322	0.323	0.538	0.335	0.419	0.384	0.461	1.000	
21. Loss of interest in sex	0.173	0.174	0.161	0.313	0.194	0.247	0.189	0.118	0.181	0.181	0.106*	0.261	0.072+	0.226	0.227	0.136	0.183	0.132	0.194	0.221	1.000

$N = 502$. All correlations significant at $P < 0.01$ except those marked as * $P < 0.05$; +ns.

6   *D. P. Carmody*

Table III. Fit indices for the alternative factor models of the BDI-II.

	$\chi 2$	df	CMIN/df	CFI	RMSEA
Model 0: Independent Factors Null Model	8992.26	231	38.93	–	0.28
Model 1: One Factor	1437.19	210	6.84	0.86	0.11
Model 2: Two-factor	1329.50	209	6.36	0.87	0.10
Model 3: Three-factor	1259.43	204	6.17	0.88	0.10

representing various ethnicities, have a common culture. They attend the same school, speak a common language at school, and share many activities in daily life, such as living on the campus or commuting from urban areas. Perhaps it is the common environmental factors that created similarities at the group level. Variations in depression scores would reflect the experiences of individual students rather than large-scale differences between ethnic groups in this urban environment.

The finding that the BDI-II is appropriate for an urban, multi-ethnic population has implications for the primary care physician, who is often the first professional contacted by individuals in distress. Screening by primary care physicians for depressive symptoms and life events may lead to interventions to assist those at risk for major depression [36], suicide [37], and long-term medical conditions [38]. The confirmation that the BDI-II has similar psychometric characteristics in multi-ethnic samples to the standardization samples allows for use of the inventory as a screening tool for signs of depression in a multi-ethnic setting. Completion of a screening inventory by patients in primary care settings may prove useful in early detection of signs of depression [39].

In summary, the results of the psychometric study of the Beck Depression Inventory – Second Edition suggest that the instrument is a reliable measure of depression in a multi-ethnic, urban environment. The descriptive characteristics of the inventory with the multiethnic sample were similar to those found in the nonclinical samples reported in the test manuals. For the BDI-II, the underlying structure is composed of three factors identified as negative attitude, performance difficulty, and somatic elements.

Acknowledgements

Financial disclosure

References

[1] Bland RC. Epidemiology of affective disorders: A review. Can J Psychiatry 1997;42:367–77.

[2] Health Canada:. A report on mental illnesses in Canada. Ottawa: Health Canada Editorial Board; 2002.

[3] Wells KB, Sturm R, Sherbourne CD, Meredith LS. Caring for depression. Cambridge, MA: Harvard University Press; 1996.

[4] Christensen O, Bundgaard S, Bech P. Prevalence of clinical (major) depression in general practice using the DSM-IV version of PRIME-MD. Int J Psychiatry Clin Prac 2001;5:49–54.

[5] Lecrubier Y. Improved ability to identify symptoms of major depressive disorder (MDD) in general practice. Int J Psychiatry Clin Prac 2001;5:S3–10.

[6] Rost K, Smith GR, Guise B, Matthews D. The deliberate misdiagnosis of major depression in primary care. Arch Fam Med 1994;3:333–7.

[7] Schulberg HC, Saul M, McClelland M, et al. Assessing depression in primary medical and psychiatric practices. Arch Gen Psychiatry 1985;42:1164–70.

[8] Bland RC. Psychiatry and the burden of mental illness. Can J Psychiatry 1998;43:801–10.

[9] Culbertson FM. Depression and gender: An international review. American Psychologist 1997;52(1):25–31.

[10] Gater R, Tansella M, Korten A, Tiemens BG, Mavreas VG, Olatawura MO. Sex differences in the prevalence and detection of depressive and anxiety disorders in general health care settings: Report from the World Health Organization collaborative study on psychological problems in general health care. Arch Gen Psychiatry 1998;55: 405–13.

[11] Jorm AF. Sex and age differences in depression: A quantitative synthesis of published research. Aust NZ J Psychiatry 1987;21:46–53.

[12] Bebbington P. The origins of sex differences in depressive disorder: Bridging the gap. Int Rev Psychiatry 1996;8(4):295–332.

[13] Plant EA, Sachs-Ericsson N. Racial and ethnic differences in depression: The roles of social support and meting basic needs. J Consult Clin Psychol 2004;72:41–52.

[14] Miech RA, Shanahan MJ. Socioeconomic status and depression over the life course. J Health Social Behav 2000;41:162–76.

[15] Twenge JM, Nolen-Hoeksma S. Age, gender, race, socioeconomic status, and birth cohort differences on the Children's Depression Inventory: a meta-analysis. J Abnorm Psychol 2002;111:578–88.

[16] Kirmayer LJ. Cultural variations in the clinical presentation of depressions and anxiety: implications for diagnosis and treatment. J Clin Psychiatry 2001;62:22–8.

[17] Oquendo MA, Ellis SP, Greenwald S, Malone KM, Weissman MM, Mann JJ. Ethnic and sex differences in suicide rates relative to major depression in the United States. Am J Psychiatry 2001;158:1652–8.

[18] Iwata N, Turner RJ, Lloyd DA. Race/ethnicity and depressive symptoms in community-dwelling young adults: a differential item functioning analysis. Psychiatry Res 2002;110:281–9.

[19] Kleinman A. Depression, somatization and the new cross-cultural psychiatry. Soc Sci Med 1977;11:3–10.

[20] Beck AT, Ward CH, Mendelson M, et al. An inventory for measuring depression. Arch Gen Psychiatry 1961;4:561–71.

[21] Beck AT, Steer RA, Brown GK. Manual for the Beck Depression Inventory, Second Edition. San Antonio, TX: The Psychological Corporation; 1996.

[22] Murphy LL, Impara JC, Plake BS. Tests in print V. Buros Institute of Mental Measurements, University of Nebraska-Lincoln, Lincoln, NE, 1999.

[23] Plake BS, Impara JC. The fourteenth mental measurements yearbook. Lincoln, NE: Buros Institute of Mental Measurements, University of Nebraska-Lincoln, 2001.

[24] Dozois DJA, Dobson KS, Ahnberg JL. A psychometric evaluation of the Beck Depression Inventory – II. Psychol Assess 1998;10:83–9.

[25] Whisman MA, Perez JE, Ramel W. Factor structure of the Beck Depression Inventory – second edition (BDI-II) in a student sample. J Clin Psychol 2000;56:545–51.

[26] Kojima M, Furukawa TA, Takahashi H, et al. Cross-cultural validation of the Beck Depression Inventory-II in Japan. Psychiatry Res 2002;110:291–9.

[27] Byrne BM, Baron P. The Beck Depression Inventory: Testing and cross-validating a hierarchical factor structure for nonclinical adolescents. J Counsel Dev 1993;26:164–78.

[28] Osman A, Downs WR, Barrios FX, et al. Factor structure and psychometric characteristics of the Beck Depression Inventory-II. J Psychopathol Behav Assess 1997;19:359–76.

[29] Buckley TC, Parker JD, Heggie J. A psychometric evaluation of the BDI-II in treatment-seeking substance abusers. J Subst Abuse Treat 2001;20:197–204.

[30] Novy DM, Nelson DV, Berry LA, Averill PM. What does the Beck Depression Inventory measure in chronic pain? A reappraisal. Pain 1995;61:261–70.

[31] Steer RA, Ball R, Ranieri WF. Further evidence for the construct validity of the Beck Depression Inventory-II with psychiatric outpatients. Psychol Rep 1997;80:443–6.

[32] World Medical Organization. Declaration of Helsinki. Br Med J 1996;313:1448–9.

[33] Arbuckle JL, Wothke W. Amos 4.0 user's guide. Chicago, IL: SmallWaters Corporation; 1999.

[34] Cole DA. Utility of confirmatory factor analysis in test validation research. J Consult Clin Psychol 1987;55:584–94.

[35] Azocar F, Arean P, Miranda J, Munoz RF. Differential item functioning in a Spanish translation of the Beck Depression Inventory. J Clin Psych 2001;57(3):355–65.

[36] Pine DS, Cohen P, Johnson JG, Brook JS. Adolescent life events as predictors of adult depression. J Affect Disord 2002;68:49–57.

[37] Schneider B, Müller M, Phillip M. Mortality in affective disorders. J Affect Disord 2001;65:263–74.

[38] Gagnon LM, Patten SB. Major depression and its association with long-term medical conditions. Can J Psychiatry 2002;47:149–52.

[39] Klinkman MS, Schwenk TL, Coyne JC. Depression in primary care – more like asthma than appendicitis: The Michigan Depression Project. Can J Psychiatry 1997;42:966–73.

AUTHOR'S QUERY SHEET

Dear Author

Some questions have arisen during the preparation of your manuscript for typesetting. Please consider each of the following points below and make any corrections required in the proofs.

Please do not give answers to the questions on *this* sheet. All corrections should be made directly in the printed proofs.