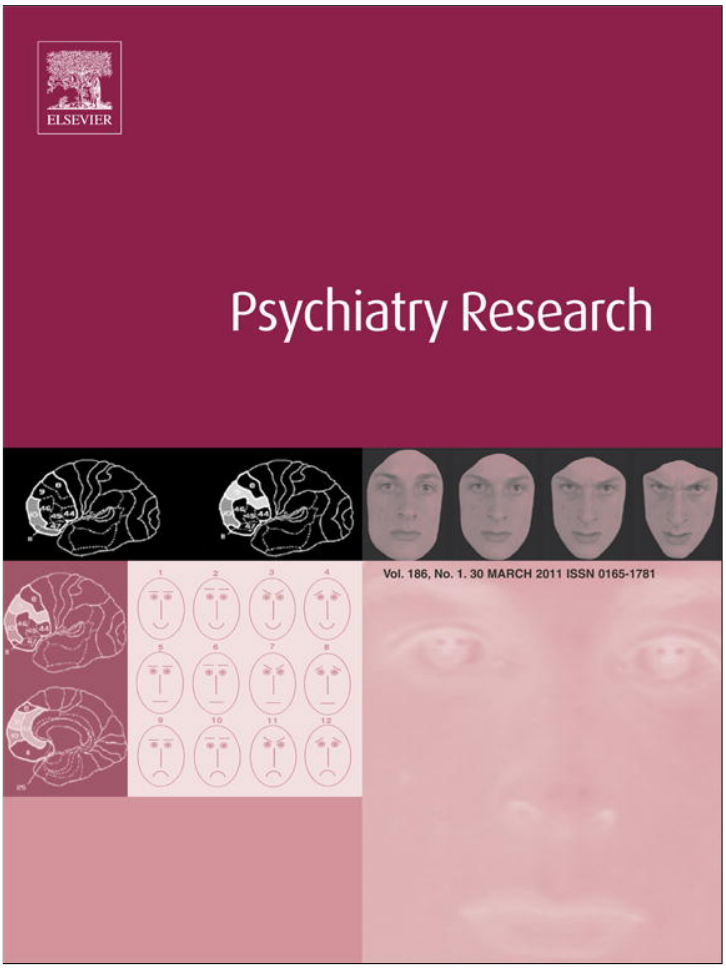

Psychiatry Research 186 (2011) 128–132

Contents lists available at ScienceDirect

Psychiatry Research

journal homepage: www.elsevier.com/locate/psychres

Validation of the Center for Epidemiologic Studies Depression Scale—Revised (CESD-R): Pragmatic depression assessment in the general population

Nicholas T. Van Dam *, Mitch Earleywine

Department of Psychology, University at Albany, SUNY, Albany, NY, USA

ARTICLE INFO

Article history:
Received 9 May 2010
Received in revised form 14 August 2010
Accepted 17 August 2010

Keywords:
Depression
CESD-R
CES-D
Assessment
Psychometrics
Diagnosis

ABSTRACT

Depression has a huge societal impact, making accurate measurement paramount. While there are several available measures, the Center for Epidemiological Studies Depression Scale (CESD) is a popular assessment tool that has wide applicability in the general population. In order to reflect modern diagnostic criteria and improve upon psychometric limitations of its predecessor, the Center for Epidemiologic Studies Depression Scale Revised (CESD-R) was recently created, but has yet to be publicized. This study explored psychometric properties of the CESD-R across a large community sample ($N=7389$) and smaller student sample ($N=245$). A newly proposed algorithmic classification method yielded base-rates of depression consistent with epidemiological results. Factor analysis suggested a unidimensional factor structure, but important utility for two separate symptom clusters. The CESD-R exhibited good psychometric properties, including high internal consistency, strong factor loadings, and theoretically consistent convergent and divergent validity with anxiety, schizotypy, and positive and negative affect. Results suggest the CESD-R is an accurate and valid measure of depression in the general population with advantages such as free distribution and an atheoretical basis.

1. Introduction

Major depression affects approximately 5% of the US population annually (Hasin et al., 2005). Although understanding of depression has increased dramatically in recent years, the most commonly used assessment scales are approximately 25 years old. A literature search in PSYCH ARTICLES suggests the most popular depression scales are the Beck Depression Inventory (BDI; Beck et al., 1961), the Center for Epidemiologic Studies Depression Scale (CES-D; Radloff, 1977), and the Hamilton Rating Scale for Depression (HAM-D; Hamilton, 1960). Psychometric flaws and clinician administration are major drawbacks to the use of the HAM-D (Bagby et al., 2004). In contrast, the BDI and CES-D are both self-administered, making them popular alternatives. Design differences (e.g., different targeted populations, response option formats, and emphasis on different aspects of depression) between the BDI and CES-D lead to measurement of different variations of depression (Skorikov and VanderVoort, 2003). Structural equation modeling analyses suggest a standardized factor loading of 0.51 and 0.63 for the BDI and CES-D, respectively, on a common hierarchical factor, suggesting that the latent traits for the BDI and CES-D are similar, but ultimately different (Skorikov and VanderVoort, 2003).

Neither the BDI nor the CES-D reflects current diagnostic criteria for major depression (American Psychiatric Association, 1994). While the revised form of the BDI (BDI-II; Beck et al., 1996) is growing in popularity, the revision of the CES-D (CESD-R; Eaton et al., 2004) is scarcely mentioned in the literature. The revision of the CES-D was undertaken to more reliably indicate general dysphoria and reflect the nine primary symptoms of a major depressive episode according to DSM-IV (Eaton et al., 2004). Among other aspects of the revision, an extra response category (nearly every day for two weeks) was added, two existing items were simplified, items predominantly unrelated to modern notions of depression were removed, and items reflecting anhedonia, psychomotor retardation/agitation, and suicidal ideation were added (see Eaton et al., 2004). In attempt to optimize the psychometric properties, items reflecting positive affect were eliminated. Some have argued that negative responses to positive items may reflect important aspects of depression (e.g., Wood et al., 2010). Several new items seem to directly reflect the inverse of old positive items. Thus, the important discriminative contribution of positive affect (e.g., Mineka et al., 1998) may potentially be compensated via better representation of dysphoria and anhedonia.

The present study aimed to determine the factor structure and to examine psychometric properties of the CESD-R across a large community sample and a smaller student sample. Comparison of base-rates of depression identified by the CESD-R in a large community sample to those identified by prior epidemiological research (e.g., Hasin et al., 2005) can

* Corresponding author. University at Albany, SUNY, Department of Psychology, Social Sciences 399, 1400 Washington Avenue, Albany, NY 12222, USA. Tel.: +1 518 445 5533; fax: +1 518 442 4867.
E-mail address: NV122142@albany.edu (N.T. Van Dam).

0165-1781/$ – see front matter © 2010 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.psychres.2010.08.018

provide important validation of the measure. Research has shown that the traditional CES-D cut-score, which also applies to the CESD-R (Eaton et al., 2004), lacks specificity (e.g., Santor et al., 1995). Therefore, an additional goal was to compare the original classification method (cut-score of 16) to an algorithmic method based on the DSM-IV approach. Correlates of the scale were also examined to explore convergent and divergent validity, as follows.

Depression and anxiety are known to be highly co-occurring conditions (Mineka et al., 1998). Accordingly, a measurement of depression should have a high correlation with a measure of anxiety. Similarly, but to a much lesser extent, schizotypy has shown positive correlations with emotional disorders, reflecting neurodevelopmental vulnerabilities for impaired cognitive and affective regulation. In particular, recent work suggests moderate relationships between positive dimensions of schizotypy and affective disorders (Lewandowski et al., 2006). Finally, while increased negative affect (NA) is common to emotional disorders (e.g., anxiety and depression), deficits in positive affect (PA) are specific to depression (Mineka et al., 1998). Since NA is common to emotional disorders, both anxiety and depression scales should significantly correlate with NA, even if common variance is controlled for (e.g., including both depression and anxiety simultaneously in the correlations). A valid measure of depression should also significantly correlate with PA after controlling for anxiety, but anxiety should not correlate with PA after controlling for depression.

2. Methods

2.1. Procedure

Participants either (a) responded to an email request issued to the National Organization for the Reformation of Marijuana Laws (NORML) listserv (Sample 1) or (b) completed questionnaires for undergraduate course credit at a state university in the northeastern United States (Sample 2). Upon completion of the survey, sample 1 participants were entered into a drawing for a $250 Amazon.com giftcard or 1 of 5 4 GB iPods. Willing participants from sample 1 forwarded the email to others who might be interested. All procedures were approved by the local institutional review board.

2.2. Participants

2.2.1. Sample 1

10,304 participants responded to the email request. Participant responses were screened for repeated Internet Protocol (IP) addresses and item completion rates less than 95%, for which 501 and 2414 cases were excluded, respectively. The final number of participants was 7389 (71.7% of original). The majority of the sample was male (80.7%) with an average age of 30.6 years (S.D. = 13.1). Participants self-identified as Caucasian (89.4%), Hispanic/Latino (5.7%), Other (1.8%), African American (1.3%), Asian (1.2%), and Native American (0.6%). The sample was relatively well educated, 44.2% reporting completion of some college, 14.8% reporting a Bachelor's degree, 13.6% reporting a High School Education, 10% reporting an Associate's degree, 8.1% reporting Some High School, 5.4% reporting an Advanced Degree, and 3.8% reporting some graduate education.

2.2.2. Sample 2

245 undergraduate psychology students participated, 62% female. The majority of participants were in their first year of university (70.1%), with 2nd (16.4%), 3rd (9.7%), and 4th year (3.7%) students represented in decreasing proportion. The average age was 19.6 years old (S.D. = 1.8) and students self-identified as Caucasian (72.8%), "Other" (9.5%), African American (8.2%), Asian (7.8%), and Hispanic/Latino (1.6%).

2.3. Measures

2.3.1. Center for Epidemiologic Studies Depression Scale—Revised

The CESD-R is an updated version of the CES-D (Radloff, 1977). The CESD-R consists of 20 items that closely reflect the DSM-IV criteria for depression (see Eaton et al., 2004). Participants were provided with instructions stating, "Below is a list of the ways you might have felt or behaved. Please select options that reflect how often you have felt this way in the past week or so" (modified from Eaton et al., 2004; p. 33). Five response options were provided: "not at all or less than 1 day", "1–2 days", "3–4 days", "5–7 days", "nearly every day for 2 weeks". Preliminary investigation suggests good psychometric properties (Eaton et al., 2004).

2.3.2. Positive and Negative Affect Schedule

Sample 2 completed the PANAS, a 20-item questionnaire designed to measure positive and negative affect (Watson et al., 1988), important latent variables in emotional disorders (e.g., Mineka et al., 1998). The questionnaire consists of 10 items that address positive affect (PA) and 10 items that address negative affect (NA). Participants endorse the extent to which single word items generally characterize how they feel on a 1 "very slightly or not at all" to 5 "extremely" response scale. The PANAS is widely used and has good psychometric properties (e.g., Watson et al., 1988; Mineka et al., 1998). The PANAS had good internal consistency (Cronbach's $\alpha = 0.823$).

2.3.3. State-Trait Inventory for Cognitive and Somatic Anxiety

The STICSA is a 21-item questionnaire designed to differentiate the cognitive and somatic components of anxiety. It was developed as an alternative to other popular measures of anxiety (e.g., State-Trait Anxiety Inventory), which have shown higher covariance with measures of depression than anxiety (Grös et al., 2007; Ree et al., 2008). The STICSA has excellent psychometric properties (Grös et al., 2007; Ree et al., 2008). The STICSA had good internal consistency in Sample 1 (Cronbach's $\alpha = 0.885$) and Sample 2 (Cronbach's $\alpha = 0.916$).

2.3.4. Schizotypal Personality Questionnaire—Brief

The SPQ-B is an abbreviated version of the full Schizotypal Personality Questionnaire. The SPQ-B consists of 22 true/false items that assess three main factors of Schizotypal Personality Disorder (Cognitive-Perceptual Deficits, Interpersonal Deficits, and Disorganization). It has good psychometric properties and significant correlations with clinical assessment (Raine and Benishay, 1995). The SPQ-B had reasonable internal consistency in Sample 1 (Cronbach's $\alpha = 0.805$) and Sample 2 (Cronbach's $\alpha = 0.835$).

2.4. Statistical methods

Unless otherwise stated, all statistical analyses were performed with SPSS 18.0.

2.4.1. Exploratory factor analysis

Exploratory factor analysis (EFA) was conducted on a random split half ($n = 3528$) of sample 1, excluding participants who did not respond to all items of the CESD-R. To minimize potential for over- or under-identification of factors, parallel analysis (e.g., Brown, 2006) and Velicer's MAP (Velicer, 1976) were computed. Parallel analysis computes randomly generated data sets to specifications and compares the obtained eigenvalues in the raw data to those obtained by chance (see O'Connor, 2000; Brown, 2006). Velicer's MAP is a step-wise process whereby components are partialed out of the correlation matrix sequentially. The step corresponding to the lowest partial squared correlation indicates the number of components (see Velicer, 1976; O'Connor, 2000). Parallel analysis using normally distributed random data generated 1000 datasets limited to the 95th percentile with principal components analysis (O'Connor, 2000). Subsequent to these analyses, maximum likelihood estimation with varimax rotation (to maximize interpretability of factors; e.g., Brown, 2006), limited to previously identified number of factors was conducted.

2.4.2. Confirmatory factor analysis

Confirmatory factor analysis (CFA) was conducted on the second random split half ($n = 3443$) of sample 1, as well as with sample 2. Both a one-factor (see Eaton et al., 2004) and a two-factor correlated model (as suggested by EFA) were explored. Participants who did not respond to all CESD-R items in samples 1 and 2 were excluded from CFA analyses. In both cases, weighted least squares estimation was used due to the ordinal nature of response options and skewed distributional properties of the CESD-R (Brown, 2006). Correlation matrices and asymptotic covariance matrices were computed using PRELIS 2.8 (Scientific Software Inc., Lincolnwood, IL). CFA was computed using LISREL 8.8 (Scientific Software Inc., Lincolnwood, IL).

Root Mean Square Error of Approximation (RMSEA) values <0.06 were considered good, while those <0.08 were considered reasonable. Comparative Fit Index (CFI) and Tucker–Lewis Index values <0.95 were considered good, while those <0.90 were considered reasonable (see Marsh et al., 2004; Brown, 2006). The Akaike Information Criterion (AIC) is a relative fit index that considers model parsimony. There is no specific AIC value that suggests "good" model fit (Brown, 2006). However, a meaningful decrease in AIC suggests an improvement in model fit (Brown, 2006).

2.4.3. Convergent and divergent validity

Anxiety and depression have high comorbidity and genetic similarities (Mineka et al., 1998). Accordingly, a correlation of the CESD-R with the STICSA was computed for both samples. Additionally, research suggests that negative affect is common to both anxiety and depression, while a negative relation to positive affect is specific to depression (Mineka et al., 1998). Relationships between the CESD-R, STICSA, and PA and NA subscales of the PANAS were computed for sample 2. The PANAS was not present in sample 1. Finally, recent work has shown relations between schizotypy and depression, though the constructs are distinct (Lewandowski et al., 2006). A correlation was computed between the CESD-R and the SPQ-B in both samples.

2.4.4. Classification schemes

A cut-score of 16 on the CES-D is indicative of a "depressive case" in clinical samples, however, it is known to produce a high rate of false positives in nonclinical samples (Santor et al., 1995). In their revision of the CES-D, Eaton et al. (2004) suggested an algorithmic classification scheme consistent with the DSM-IV (American Psychiatric Association, 1994). To test the differential classification of the traditional

cut-score relative to an algorithmic classification for probable depression (see supplemental material), both were computed and compared.

Given the psychometric differences between clinical and nonclinical populations on the CES-D (e.g., Santor et al., 1995), factor intercorrelations were also computed for the two-factor model in the subset of sample 1 that met algorithmic classification for probable depression.

3. Results

3.1. CESD-R scale properties

In sample 1, 6971 participants completed all CESD-R items for a total score range of 0 to 77 ($M = 10.3$, S.D. $= 11.7$). The distribution of CESD-R scores had a large positive skew (2.15 ± 0.03) and was leptokurtic (kurtosis $= 5.77 \pm 0.06$). Internal consistency was high (Cronbach's $\alpha = 0.923$). In sample 2, 243 participants completed all CESD-R items for a total score range of 0 to 71 ($M = 16.4$, S.D. $= 13.5$). The distribution of CESD-R scores had a positive skew (1.49 ± 0.16) and was leptokurtic (kurtosis $= 2.56 \pm 0.32$). Internal consistency was high (Cronbach's $\alpha = 0.928$).

3.2. Exploratory factor analysis

Parallel analysis suggested four roots with eigenvalues larger than what would be obtained by chance. Velicer's MAP revealed a smallest average squared partial correlation of 0.020 on step two suggesting two underlying components. Maximum Likelihood estimation using promax rotation limited to two-, three-, and four-factor solutions was used to explore factor loadings. Examination of the three- and four-factor solutions revealed inconsistencies with theoretical considerations and optimal psychometric properties (see Brown, 2006). Both solutions suggested a factor containing only three items (1, 5, 19) related to appetite and sleep. This factor excluded another item related to sleep (11) and one related to weight changes (18), suggesting substantive inconsistencies. The two-factor solution was both theoretically and psychometrically consistent, suggesting one factor related to negative mood and another factor related to functional impairment. See Table 1.

3.3. Confirmatory factor analysis

Fit indices are displayed in Table 2. All fit indices for sample 1 were indicative of good model fit. In sample 1, a two-factor model fit the data significantly better than a one-factor solution, $\Delta\chi^2_{(1)} = 124.7$, $p < 0.001$. This was also suggested by a decrease of 122.7 in AIC. Two of three indices for sample 2 were indicative of good model fit, while the RMSEA was larger than values considered reasonable. It should be noted that RMSEA values can be impacted by sample size necessitating multiple evaluation methods (see Brown, 2006). In sample 2, a two-factor model again fit the data significantly better than a one-factor solution, $\Delta\chi^2_{(1)} = 31.1$, $p < 0.001$. This was again suggested by a decrease of 29.1 in AIC. Despite the suggestion that a two-factor model was an improvement over the one-factor model, the inter-factor correlation was 0.941 in sample 1 and 0.975 in sample 2, suggesting factor redundancy (Brown, 2006). This suggests a one-factor model is most parsimonious.

Table 1
Standardized factor loadings from confirmatory factor analysis for second half of sample 1 ($N = 3650$).

Item	1 Factor	2 Factor	
	D	FI	NM
1. My appetite was poor.	0.646	0.646	–
2. I could not shake off the blues.	0.953	–	0.950
3. I had trouble keeping my mind on what I was doing.	0.884	0.874	–
4. I felt depressed.	0.968	–	0.967
5. My sleep was restless.	0.907	0.901	–
6. I felt sad.	0.956	–	0.954
7. I could not get going.	0.907	0.896	–
8. Nothing made me happy.	0.975	–	0.971
9. I felt like a bad person.	0.945	–	0.940
10. I lost interest in my usual activities.	0.935	0.926	–
11. I slept much more than usual.	0.621	0.622	–
12. I felt like I was moving too slowly.	0.900	0.887	–
13. I felt fidgety.	0.757	0.747	–
14. I wished I were dead.	0.959	–	0.961
15. I wanted to hurt myself.	0.945	–	0.942
16. I was tired all the time.	0.869	0.864	–
17. I did not like myself.	0.963	–	0.961
18. I lost a lot of weight without trying to.	0.614	0.599	–
19. I had a lot of trouble getting to sleep.	0.930	0.930	–
20. I could not focus on the important things.	0.939	0.931	–

D = Depression, FI = Functional Impairment, NM = Negative Mood.

3.4. Convergent and divergent validity

3.4.1. Sample 1

There was a large positive correlation between the CESD-R and the STICSA, $r = 0.737$, $p < 0.01$. There was also a medium positive correlation between the CESD-R and the SPQ-B, $r = 0.436$, $p < 0.01$. Sample 1 did not complete the PANAS.

3.4.2. Sample 2

There was a large positive correlation between the CESD-R and STICSA, $r = 0.653$, $p < 0.01$. There was also a medium positive correlation between the CESD-R and SPQ-B, $r = 0.426$, $p < 0.01$. The PANAS-NA was positively correlated with both the CESD-R, $r = 0.576$, $p < 0.01$, and STICSA, $r = 0.663$, $p < 0.01$. The PANAS-PA was negatively correlated with both the CESD-R, $r = -0.263$, $p < 0.01$, and STICSA, $r = -0.233$, $p < 0.01$. In order to explore the specificity of PA to depression, correlations of PA with anxiety, controlling for depression, and depression, controlling for anxiety, were computed. When controlling for CESD-R, there was no relation between PA and STICSA, $\beta = -0.041$, $t = -0.74$, $p = 0.46$, suggesting that the included covariance of depression removed significant relations between PA and anxiety. The same did not hold when the STICSA was controlled between PA and CESD-R, $\beta = -0.160$, $t = -3.00$, $p = 0.003$, suggesting that the included covariance of anxiety did not alter significant relations between PA and depression.

3.5. Classification schemes

The proportion of individuals in sample 1 meeting algorithmic classification criteria was 4.6%, almost the exact percentage of individuals classified as depressed in the largest epidemiological study of depression to date (95% confidence interval = 4.98–5.57; Hasin et al., 2005). In contrast, the 16+ method classified 16.5% of the sample as depressive cases, corresponding to 75 standard error values above the mean reported by Hasin and colleagues. Comparison of the two classification schemes shows that relative to the algorithmic classification, the 16+ method "falsely" identifies 11.9% of cases as

Table 2
Fit indices for confirmatory factor analyses.

	Model	χ^2	d.f.	AIC	CFI	RMSEA	TLI
Sample 1[a]	1 Factor	1396.2	170	1476.2	0.965	0.046	0.961
	2 Factor	1271.5	169	1353.5	0.969	0.044	0.965
Sample 2[b]	1 Factor	819.2	170	899.2	0.995	0.126	0.994
	2 Factor	788.1	169	870.1	0.997	0.123	0.995

χ^2 = chi-square statistic, d.f. = degrees of freedom, AIC = Akaike Information Criterion, CFI = Comparative Fit Index.
RMSEA = Root Mean Square Error of Approximation, TLI = Tucker–Lewis Index.
[a] Sample 1 comprises the second random split half of a sample invited to participate via email ($n = 3443$).
[b] Sample 2 comprises 243 undergraduate psychology students who received course credit for participation.

probably depressed. It is important to note that the algorithmic classification scheme does not identify unique cases, just a smaller subset of those identified by the cut-score. In the "probably depressed" sample, those who met criteria according to the algorithm, the correlation between the two factors was 0.499, suggesting the potential import of the separate symptom clusters of negative mood and functional impairment in depressive cases.

4. Discussion

4.1. General discussion

Exploratory and confirmatory factor analyses, assessment of internal consistency, and exploration of convergent and divergent validity all suggest the CESD-R has strong psychometric properties, making it a useful tool for assessing depression in the general population. The CESD-R also has theoretically consistent relations to measures of anxiety, negative affect, and positive affect (e.g., Mineka et al., 1998). These findings are particularly salient given that the majority gender is opposite in the two samples. Additional evidence is derived from the algorithmic classification scheme, which identified a base-rate of depression comparable to the largest epidemiological study of depression to date (Hasin et al., 2005). The algorithmic classification scheme is consistent with the DSM-IV diagnostic content and process. This is particularly salient given that criteria for a major depressive episode are not likely to change with the introduction of DSM-V. The algorithmic approach maintains the sensitivity of the cut-score method while seemingly improving the specificity of identifying depressive cases.

The CESD-R may additionally be the able to evaluate depressive severity. Endorsement of symptom frequency would seem to lend itself to dimensional classification (Kraemer et al., 2004). Research has shown that functional impairment, one aspect of what the CESD-R seems to measure, is an important indicator of depressive severity (Mintz et al., 1992). Additionally, correlations between CESD-R total score and convergent measures suggest that total score indicates depressive severity. Researchers might consider a dichotomous classification based on the algorithmic classification scheme (depressed or not depressed), supplemented with a dimensional score indicating severity.

While extremely high inter-factor correlations for the CESD-R suggest one latent trait, this might vary in a clinical sample. When considering only those individuals classified by algorithmic criteria as "probably depressed", the inter-factor correlation decreases to 0.499, indicating that 75% of variance in the two factors is not overlapping. This suggests two symptom clusters, negative mood and functional impairment. Negative mood likely represents dysfunctional attitudes, beliefs, and/or schemata that underlie cognitive risk factors for depression while functional impairment likely represents the extent to which negative mood is interfering with daily life (e.g., Mintz et al., 1992). The two symptom clusters may represent different aspects of depression, with combinations representing different depressive phenotypes. The differential presentation of depression across cultural and ethnic groups further enhances the import of these two clusters. Ethnic minorities are more likely to experience functional impairment and to report more somatic symptoms than majority counterparts (e.g., Plant and Sachs-Ericsson, 2004). Thus, functional impairment may represent an important means of measuring alternative manifestations of depression while negative affect more traditional components. All of these findings suggest the CESD-R is a useful tool for assessing depression in the general population.

4.2. Limitations

Sample 1 participants were targeted via a listserv related to marijuana policy, which makes the population more likely to use marijuana and other drugs (e.g., Looby and Earleywine, 2007). Relations between cannabis use and mood disorder incidence warrant consideration, though longitudinal research negates a causal role of cannabis use in depression (Harder et al., 2006). In a comparable sample where 32% of individuals were daily cannabis users, only 12% met proxy criteria for dependence (Looby and Earleywine, 2007). While individuals with dependence had higher CES-D scores, only about 39% of those individuals were at or above the original cut-score of 16, about 1.5% of the total sample. Regardless, we cannot be certain that drug use did not influence response styles on questionnaires in the current study.

Additionally, both samples were predominantly Caucasian and well educated, limiting generalization to more diverse groups. One of the major strengths of the CES-D is its validation across numerous populations (Eaton et al., 2004). The CESD-R has been translated into Chinese, French, Japanese, and Spanish (personal communication, W. Eaton, July 29, 2010), and the Spanish version has shown excellent reliability and good item-total correlations (Eaton et al., 2004). Nevertheless, further exploration of the factor structure and validity of the CESD-R in more diverse populations is warranted.

Another limitation is the lack of additional verification of depression. Future work should include additional measures of depression, particularly the CES-D. While previous comparisons to the CES-D have shown high correlations (Eaton et al., 2004), a comparison of criterion validity would prove useful. While the CESD-R was designed for the general public, examination in a clinical sample would be beneficial. Eaton et al. (2004) validated the CESD-R in a small clinical sample and over 2000 participants from a more general population. Future examination of psychometric properties in clinical samples would permit examination of the two proposed symptom clusters.

4.3. Conclusions

The CESD-R is an accurate and valid measure of depression that is freely available and atheoretical. The present study suggests a single factor with important implications for two main symptom clusters (negative mood and functional impairment). The two clusters represent salient aspects of depression that may differ across phenotypes of depression. The algorithmic classification method provides an efficient and pragmatic approach to identifying probable depression, improving accuracy while seemingly maintaining the sensitivity of the original CES-D. The algorithmic approach could be supplemented with dimensional scores on the two symptom clusters to index depressive severity and phenotype. Despite potential utility as an indicator of probable depression, the present analyses should not be taken to indicate that a self-report measure is a valid means of diagnosis. Clearly, accurate diagnosis requires clinical skill and expertise, though an accurate measure of depressive symptomatology might aid such diagnosis. It seems that the CESD-R may maintain advantages of the original scale, while providing new benefits.

Acknowledgements

The authors thank Dick Haase and Gerhard Mels for technical advice and Gillinder Bedi and Kyle DeYoung for comments on earlier drafts.

Appendix A. Supplementary data

Supplementary data to this article can be found online at doi: 10.1016/j.psychres.2010.08.018.

References

American Psychiatric Association, 1994. Diagnostic and Statistical Manual of Mental Disorders, 4th ed. American Psychiatric Association, Washington, DC.

Bagby, R.M., Ryder, A.G., Schuller, D.R., Marshall, M.B., 2004. The Hamilton Depression Rating Scale: has the gold standard become a lead weight? American Journal of Psychiatry 161, 2163–2177.
Beck, A.T., Steer, R.A., Brown, G.K., 1996. Manual for the Beck Depression Inventory-II. Psychological Corporation, San Antonio, TX.
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.
Brown, T.A., 2006. Confirmatory Factor Analysis for Applied Research. Guilford Press, New York, NY.
Eaton, W.W., Smith, C., Ybarra, M., Muntaner, C., Tien, A., 2004. Center for Epidemiologic Studies Depression Scale: review and revision (CESD and CESD-R), In: Maruish, M.E. (Ed.), The Use of Psychological Testing for Treatment Planning and Outcomes Assessment, 3rd ed. : Instruments for Adults, 3. Lawrence Erlbaum, Mahwah, NJ, pp. 363–377.
Grös, D.F., Antony, M.M., Simms, L.J., McCabe, R.E., 2007. Psychometric properties of the State-Trait Inventory for Cognitive and Somatic Anxiety (STICSA): comparison to the State-Trait Anxiety Inventory (STAI). Psychological Assessment 19, 369–381.
Hamilton, M., 1960. A rating scale for depression. Journal of Neurology, Neurosurgery, and Psychiatry 23, 56–62.
Harder, V.S., Morral, A.R., Arkes, J., 2006. Marijuana use and depression among adults: testing for causal associations. Addiction 101, 1463–1472.
Hasin, D.S., Goodwin, R.D., Stinson, F.S., Grant, B.F., 2005. Epidemiology of Major Depressive Disorder: results from the National Epidemiologic Survey on Alcoholism and Related Conditions. Archives of General Psychiatry 62, 1097–1106.
Kraemer, H.C., Noda, A., O'Hara, R., 2004. Categorical versus dimensional approaches to diagnosis: methodological challenges. Journal of Psychiatric Research 38, 17–25.
Lewandowski, K.E., Barrantes-Vidal, N., Nelson-Gray, R.O., Clancy, C., Kepley, H.O., Kwapil, T.R., 2006. Anxiety and depression symptoms in psychometrically identified schizotypy. Schizophrenia Research 83, 225–235.
Looby, A., Earleywine, M., 2007. Negative consequences associated with dependence in daily cannabis users. Substance Abuse: Treatment, Prevention, and Policy 2, 3.
Marsh, H.W., Hau, K.T., Wen, Z., 2004. In search of golden rules: comment on hypothesis-testing approaches to setting cut-off values for fit indices and dangers to overgeneralizing Hu and Bentler's (1999) findings. Structural Equation Modeling 11, 320–341.
Mineka, S., Watson, D., Clark, L.A., 1998. Comorbidity of anxiety and unipolar mooddisorders. Annual Review of Psychology 49, 377–412.
Mintz, J., Mintz, L.I., Arruda, M.J., Hwang, S.S., 1992. Treatments of depression and the functional capacity to work. Archives of General Psychiatry 49, 761–768.
O'Connor, B.P., 2000. SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. Behavior Research Methods, Instrumentation, and Computers 32, 396–402.
Plant, E.A., Sachs-Ericsson, N., 2004. Racial and ethnic differences in depression: the roles of social support and meeting basic needs. Journal of Consulting and Clinical Psychology 72, 41–52.
Radloff, L.S., 1977. The CES-D scale: a self-report depression scale for research in the general population. Journal of Applied Psychological Measurement 1, 385–401.
Raine, A., Benishay, D., 1995. The SPQ-B: a brief screening instrument for schizotypal personality disorder. Journal of Personality Disorders 9, 346–355.
Ree, M.J., French, D., MacLeod, C., Locke, V., 2008. Distinguishing cognitive and somatic dimensions of state and trait anxiety: development and validation of the State-Trait Inventory for Cognitive and Somatic Anxiety (STICSA). Behavioural and Cognitive Psychotherapy 36, 313–332.
Santor, D.A., Zuroff, D.C., Cervantes, P., Palacios, J., Ramsey, J.O., 1995. Examining scale discriminability in the BDI and CES-D as a function of depressive severity. Psychological Assessment 7, 131–139.
Skorikov, V.B., VanderVoort, D.J., 2003. Relationships between the underlying constructs of the Beck Depression Inventory and the Center for Epidemiological Studies Depression Scale. Educational and Psychological Measurement 63, 319–335.
Velicer, W.F., 1976. Determining the number of components from the matrix of partial correlations. Psychometrika 41, 321–327.
Watson, D., Clark, L.A., Tellegen, A., 1988. Development of and validiation of brief measures of positive and negative affect: the PANAS scale. Journal of Personality and Social Psychology 54, 1063–1070.
Wood, A.M., Taylor, P.T., Joseph, S., 2010. Does the CES-D measure a continuum from depression to happiness? Comparing substantive and artifactual models. Psychiatry Research 177, 120–123.