HHS Public Access

Author manuscript

J Affect Disord. Author manuscript; available in PMC 2015 April 01.

Published in final edited form as:

J Affect Disord. 2014 April ; 158: 154–160. doi:10.1016/j.jad.2014.02.009.

Psychometric evaluation of a self-report scale to measure adolescent depression: the CESDR-10 in two national adolescent samples in the United States

Emily E. Haroz, M.A.[a,*], **Michele Ybarra, PhD**[a,b], and **William W. Eaton, PhD**[a]

[a]Johns Hopkins Bloomberg School of Public Health, Department of Mental Health, Baltimore, MD

[b]Center for Innovative Public Health Research, San Clemente, CA

Abstract

Background—There is a need for brief instruments to screen for depression in adolescents that are valid, reliable and freely available. The aim of this study was to investigate the psychometric properties of a 10-item version of the CESD-R (CESDR-10) in two national adolescent samples.

Methods—Sample 1 consisted of *N*=3777 youths (mean age 15.7) and Sample 2 contained *N*=1150 adolescents (mean age 14.5). We performed confirmatory factor analysis, evaluated construct validity, examined differential item functioning, and assessed internal consistency reliability (α).

Results—The results suggest generally strong psychometric properties for the CESDR-10. The CFA 1-factor model showed good model fit. Construct validity was partially supported in Sample 1 and mostly supported for Sample 2 based upon the characteristics examined. The CESDR-10 showed configural and metric invariance across both samples and full measurement invariance across sex. There were no notable differences in discrimination parameters or clinically significant differential item functioning between samples or sexes.

Limitations—Criterion related validity was not assessed in this study. Further studies should evaluate the scale in comparison to a psychiatric diagnosis. In addition, this study utilized a web-based format of administration which may influence participants' answers. In future studies, the CESDR-10 should be administered in other settings to more thoroughly establish its generalizability.

[*]Corresponding author: eharoz@jhsph.edu, Johns Hopkins Bloomberg School of Public Health, Department of Mental Health, 624 N. Broadway, Room 780 Baltimore, MD 21205.

Contributors:

Emily E. Haroz – Performed all statistical analysis, wrote the manuscript

Michele Ybarra – Provided the data, provided oversight and guidance on statistical analysis, contributed significantly to the manuscript writing process

William Eaton – Contributed significantly to the manuscript, provided oversight and guidance on statistical analysis

Conflict of Interest: None declared

Conclusion—In clinical and non-clinical settings alike, time pressures make the availability of brief but valid screening measures critical. Findings support future use of the CESDR-10.

Keywords

Depression; Adolescence; Assessment; CESD-R; Psychometrics

Introduction

With a 12-month prevalence rate of 4–5% (Jane Costello, Erkanli et al. 2006, Thapar, Collishaw et al. 2012), depression is relatively common in adolescence (Murray, Jamison et al. 2006, Kilpatrick, Ruggiero et al. 2003). Depression is a risk factor for adolescent suicide, and leads to increased risk of substance use, obesity, social difficulties, and educational problems (Thapar, Collishaw et al. 2012, Keenan-Miller, Hammen et al. 2007, Fletcher 2010). While many scales exist for the measurement of adult depressive symptomatology (Beck, Ward et al. 1961, Derogatis, Lipman et al. 1974, Hamilton 1960, Kroenke, Spitzer et al. 2001, Radloff 1977), fewer well-validated scales exist that measure depressive symptoms specifically with adolescents. Age-appropriate screeners are needed because the symptom list varies slightly for adolescents and includes an additional symptom, irritability (American Psychiatric Association 2000, Stringaris, Zavos et al. 2012).

Brief measures such as the two-item patient health questionnaire (PHQ-2) (Kroenke, Spitzer et al. 2003) and the Strengths and Difficulties Questionnaire (Goodman 1997), have been shown to be adequate screening tools for adolescents, yet these scales do not reflect current diagnostic criteria. There are several longer scales that are well validated, such as the Reynolds Adolescent Depression Scale (30-items) (Reynolds 1986), the Short Mood and Feelings Questionnaire (Messer, Angold et al. 1995) (13-items) and the Children's Depression Inventory-2 (25-items) (Kovacs 2005) but these often require lengthy administration, making them difficult to use in time constrained settings such as schools, health clinics, and in survey research. Both the RADS and the CDI exist in 10-item versions, but the shorter versions may lack specificity (Allgaier, Frühe et al. 2012, Milfont, Merry et al. 2008) and are not freely available. Moreover, the symptom of irritability is often not included in standard depression questionnaires, but is an important symptom to evaluate in adolescents (Pickles, Aglan et al. 2010). Certainly too, scales that are not freely available (Kovacs 2005, Reynolds 1986) are not as feasible to use in non-research settings. There continues to be a need for instruments that reflect the range of depressive symptoms adolescents experience; that are valid, reliable and freely available; and that can be used in a variety of epidemiological and clinical settings.

Background

The CES-D was developed in the 1970s to assess depression symptoms in community-based samples and epidemiological surveys (Radloff 1977). The CES-D has been widely used and translated into several different languages. It also has been used and tested with adolescents (Radloff 1991). In 2004, Eaton et al. revised the original CES-D to reflect the criteria in the DSM-IV-TR (American Psychiatric Association 2000). The revision resulted in a 20-item

instrument (CESD-R) that demonstrated strong psychometric properties (Eaton, Smith et al. 2004).

The CESD-R has now been independently evaluated (Van Dam, Earleywine 2011) and is freely available online (Eaton, Ybarra et al. 2012). While others have shortened the original CES-D to various 10-item (or less) versions (Andresen, Malmgren et al. 1994, Cole, Rabin et al. 2004, Irwin, Artin et al. 1999, Kohout, Berkman et al. 1993, Santor, Coyne 1997) including a 10-item version with adolescents (Cartierre, Coulon et al. 2011), the original CES-D was not developed based on DSM (American Psychiatric Association 2000) criteria. In fact, all of the shorter versions mentioned above only include between 2 and 5 of the 9 DSM-IV symptoms for a major depressive episode (5 DSM-IV symptoms (Kohout, Berkman et al. 1993); 4 DSM-IV symptoms (Andresen, Malmgren et al. 1994, Cheng, Chan et al. 2006, Cole, Rabin et al. 2004, Irwin, Artin et al. 1999, Santor, Coyne 1997, Kohout, Berkman et al. 1993); 2 DSM-IV symptoms (Burnam, Wells et al. 1988)). As such, these brief scales may reflect outdated and under-realized conceptualizations of depressive disorder. Revising the CESD-R to a shorter scale that accurately and comprehensively reflects the presentation of adolescent depression has the potential to enhance screening by better capturing symptoms of adolescent depression. To date, a brief adolescent version based upon the CESD-R is lacking.

Current Study

The aim of this study was to create and evaluate a 10-item version of the CESD-R, which would have the advantage of brevity, the possibility of linking directly to the CESD-R (in nine of its items), and which would include the symptom of irritability making it advantageous for use with adolescents. Results from this investigation may support the use of this scale to assess adult as well as adolescent depression symptoms in community settings in the United States.

Methods

Participants

Study samples were drawn from two national surveys involving adolescents: the Teen Health and Technology study (THT; Sample 1) and the Growing Up with Media project (GuwM; Sample 2). The THT study was a cross-sectional study of *N*=5680, 13–18 year old male, female and transgendered youth. Participants were recruited nationally either: randomly through the Harris Poll OnLine (HPOL), or through targeted outreach efforts by the Gay Lesbian Straight Education Network. To be eligible, youth had to be U.S. residents ages 13 to 18 and be in the 5^{th} grade or above. Because the aim of the current paper was to validate the measure in the general population, only the randomly recruited HPOL sample was included (*n*=3777). Each participant was weighted based on sex, race/ethnicity, parents' level of education, school location and geographical region. All participants provided completed informed assent to participate and surveys were completed online.

Similarly, the GuwM study recruited parent-child pairs nationally through the HPOL. Households were randomly identified and recruited through adult HPOL members. To be

eligible for the GuwM study, adults had to be a parent/guardian of a 10–15 year old child who lived in their household at least 50% of the time and be familiar with the child's daily activities. The GuwM survey was conducted yearly over a three-year period. Depression items were added at Wave 3. As such, the current analyses focuses on data from *n*=1150 youths who responded at this wave. The final GuwM data were weighted based on age, biological sex, race/ethnicity, region, education, and household income. Additional weights were included that accounted for propensity to be online, join online panels, respond to this particular survey, and participate in this survey after Wave1. Caregivers provided informed consent and permission for their child's participation. Youth provided informed assent to participate. All consent and surveys were completed online.

Measures

Measures consisted of the scale under investigation, the CESDR-10, and additional scales used to assess construct validity.

Center for Epidemiologic Studies Depression Scale-Revised 10-Item Version for Adolescents (CESDR-10)—Adolescent depressive symptoms in both samples were assessed using an adapted version of the CESD-R (Eaton, Smith et al. 2004). Items for this 10-item adolescent-specific version were selected based on the highest factor loadings form each of the 9 symptom domains in the CESD-R (Eaton, Smith et al. 2004).. Each item is scored on a 5-point ordinal scale for frequency (*0=not at all or less than 1 day in the last 2 weeks*; *1=1–2 days in the last week*; *2=3–4 days in the last week*; *3=5–7 days in the last week*; *4=nearly every day for 2 weeks*). The THT study used a 0–4 response range, while the GuwM study used a 1–5 response range.

Eaton, Ybarra & Schwab (Eaton, Ybarra et al. 2012) have proposed a scoring algorithm for the CESD-R to determine possible depressive symptom categories. For the CESDR-10, this algorithm was revised to fit DSM criteria for depressive disorders. For scoring purposes in terms of the subthreshold depression and no clinical significance categories, the response categories were revised to match the scale of the CESD-R. This meant collapsing the categories *5–7 days* and *nearly every day for 2 weeks* into one category. For the other scoring categories the 5-point scales was retained.

Compared to the CESD-R (20-items), CESDR-10 scoring was based on endorsement of each item. Criteria for Major depressive episode was defined as 1) the presence of anhedonia, dysphoria, or irritability nearly every day for the past two weeks and 2) at least 4 additional symptoms endorsed as occurring nearly every day for the past two weeks. Probable major depressive episode was defined as 1) the presence of anhedonia, dysphoria or irritability nearly every day for the past two weeks, and 2) an additional 3 symptoms endorsed as occurring nearly every day for the past two weeks, or 5–7 days in the past week. Possible major depressive episode was defined as 1) the presence of anhedonia, dysphroria or irritability nearly every day for the past two weeks, and 2) an additional 2 symptoms endorsed as occurring nearly every day for the past two weeks, or 5–7 days in the past week. Finally subthreshold depression symptoms were defined as those who had a score of at least 8, but do not meet above criteria; and no clinical significance was defined as people who had

a total score of less than 8. Scoring of the CESDR-10 was performed in order to give estimates of the epidemiology of depression in the two study samples.

In addition, we evaluated the relationship between the CESDR-10 and scales for self-esteem (Rosenberg 1989), social support (Zimet, Dahlem et al. 1988), parent-child relationships (Finkelhor, Mitchell et al. 2000), substance use (Brener, Collins et al. 1995), and aggressive behavior (Dahlberg, Toal et al. 2005, Bachman, Johnston et al. 2001, Udry 1996). Research shows that both self-esteem and social support are negatively correlated with depression symptoms (Orth, Robins et al. 2008, Lin, Tang et al. 2008, Ellis, Nixon et al. 2009). Depression and substance use are thought to be highly comorbid and positively correlated (Keenan-Miller, Hammen et al. 2007). Aggressive behavior can also be related to depression in adolescents (Price, Salekin et al. 2012, Garber, Quiggle et al. 1991). Details about these measures are available upon request.

Statistical Analysis

First, scoring was performed to determine the number of participants in each survey who met criteria for a possible depressive symptom category. Next, analyses were performed to examine the scale's reliability and validity. Tests consisted of evaluation of internal consistency reliability (α), confirmatory factor analysis (CFA), assessments of construct validity, an exploration of differential item functioning. All statistical analyses were performed using STATA 11 (StataCorp 2009) and Mplus 7.1 (Muthén, Muthén 2012).

Internal consistency reliability—Internal consistency reliability provides a measure of the degree of homogeneity of the items on the scale as items on the scale should be correlated with each other and the entire score. Cronbach's alpha (α) was assessed for the CESDR-10 in both samples.

Confirmatory factor analysis—A confirmatory factor analysis (CFA) specifying one latent factor, 'depression', was performed on the CESDR-10 in both samples independently and combined. Model fit was evaluated by examining the root mean square error of approximation (RMSEA), the comparative fit index (CFI), the Tucker-Lewis Index (TLI) and the Weighted Root Mean Square Residual (WRMR). RMSEA values lower than 0.05, TLI/CFI values above 0.90, and WRMR values less than 0.90 all are indicative of good model fit (Yu, Muthen 2002).

Construct validity—Construct validity was assessed using three methods: 1)evaluatation of the relationship of the CESDR-10 to substance use, self-esteem, aggressive behavior and social support scales, using. Pearson's correlations (*r*) with sample weights; 2) examination of average CESDR-10 scores by sex and age, with the notion that there should be a divergence in average scores of depression as adolescents get older supporting a developmental emergence of differences in depression rates between sexes; and 3) examination of construct validity across studies (GuwM and THT) and sexes utilizing measurement invariance models (Park, Gross et al. 2012, Steinmetz, Schmidt et al. 2009), testing configural, metric and scalar invariance. *Configural* invariance tests if the same set of factors are present and each group. *Metric* invariance tests if factor loadings are the same

across groups. *Scalar* invariance tests if item intercepts are the same, which indicates whether there are systematic differences in group responses. To compare models, we used chi-squared difference tests ($\Delta\chi 2$). Non-significant ($p < 0.05$) chi-squared difference tests indicate that there is no difference between the models and the more constrained model may be accepted.

Item Response and Differential Item Functioning Analysis—We utilized a Samejima's graded response model (Samejima 1997) to examine item discrimination parameters across studies and sexes. This involved using weighted least-squares estimation to examine discrimination parameters within each sample and by sex. We then utilized a Multiple Indicators, Multiple Causes Model (MIMIC) using full maximum likelihood estimation on the observed frequencies to determine whether any differential item functioning (DIF) by sample or sex was present. MIMIC models have been used to detect DIF in various academic and psychological testing situations (Gallo, Anthony et al. 1994, Nuevo, Dunn et al. 2009, Finch 2005). DIF is present when individuals from two different groups, but with the same underlying level of the latent trait, have different probabilities of endorsing an item based on their group membership.

Results

Sample Characteristics

Sample 1—The THT sample included *N*=3777 youths with a mean age of 15.7 years (*SD*=1.7). Overall, there were 1639 males and 2138 females. The sample was predominantly white (75%, *n*=2841). Most participants were in high school at the time of the survey (72%) (Table I).

Sample 2—The GuwM sample consisted of *N*=1150 youths with an average age of 14.5 years (*SD*=1.8). There were 582 males and 568 females. As with the THT sample, GuwM participants were mostly white (74%, *n*=854) and the majority were currently in high school at the time of the survey (62.7%) (Table I).

Scoring of CESDR-10

In the THT sample, *n*=121 (3.2%) met criteria for major depressive episode, among whom *n*=81 were female and *n*=40 were male. For probable major depressive episode, an additional *n*=81 (2.1%) adolescents (*n*=58 females; *n*=23 males) met criteria; and for possible major depressive episode, an additional *n*=38 (1.0%) adolescents (*n*=24 females; *n*=14 males) met criteria. A total of *n*=750 (19.9%) adolescents (*n*=507 females; *n*=243 males) were determined to have subclinical depressive symptoms currently (i.e., a total of 8 or above on the CESDR-10) and *n*=2787 (73.8%), *n*=1470 females; *n*=1322 males, were classified as not having symptoms of clinical significance.

In the GuwM sample, *n*=21 (1.8%) met criteria for major depressive episode, which included *n*=16 females and *n*=5 males. An additional *n*= 23 (2.0%) met criteria for probable major depressive episode (*n*=13 females; *n*=10 males), and *n*= 11 (0.9%) (*n*=8 females; *n*=2 males) more adolescents met criteria for possible major depressive episode. A total of *n*=190

(16.5%) (*n*=112 females; *n*=78 males) of adolescents had subclinical depressive symptoms and *n*=905 (78.7%) (*n*=423 females; *n*=489 males) were classified as not having clinically significant symptoms.

Internal consistency reliability

Cronbach's alpha (α) was high in both samples; α=0.91 for the THT sample and α=0.90 for the GuwM sample. Individual level item analysis of the 10 items on the CESDR-10 in both samples indicated that removal of any of the items would have reduced the overall alphas slightly.

Confirmatory factor analysis

The CFA for the THT sample indicated that the one factor model showed good model fit (RMSEA=0.06; TLI=0.99; CFI=0.99; WRMR=1.5; see Table II). Model fit indices were not as strong for the GuwM sample (RMSEA=0.08; TLI=0.96; CFI=0.97; WRMR=1.3). For the combined samples, model fit indices were adequate (RMSEA=0.06; TLI=0.98; CFI=0.99; WRMR=1.9).

Construct validity

In the THT study, final scores on the CESDR-10 were positively correlated with total scores on the substance use scale (*r*=0.09). Correlations of specific types of substances ranged from *r*=0.10 with cigarettes to *r*=0.13 for marijuana. Total CESDR-10 scores were negatively correlated with total scores for the self-esteem measure (*r*=−0.56). The correlation between total depression and total social support scale was *r*=0.08. For in-person social support specifically, it was *r*=−0.15 and for online social support it was *r*=0.14. CESDR-10 scores were positively correlated with negative parent-child relationships (*r*=0. 35).

For the GuwM study, total scores on the CESDR-10 were positively correlated with scores on the substance use questionnaire (*r*=0.19). Correlations of specific types of substances ranged between *r*=0.19 for marijuana to *r*=0.21 for cigarettes use. Total CESDR-10 scores were positively correlated with aggressive behaviors (*r*=0.44) and negatively correlated with total social support (*r*=−0.06). Similar correlations were observed when the 'friend'(*r*=−0.06) and 'special person' (*r*=−0.07) subscales were assessed separately. The CESDR-10 was positively associated with negative parent-child relationships (*r*=0.42).

Across all ages in both samples, female adolescents had higher average scores on the CESDR-10 than their male counterparts. In the GuwM sample, differences between sexes begin to emerge for the 13–14 year-old age group: Male adolescents scored an average of 13.6 and females 16.8. Similar differences were observed for this age group (13–14) in the THT sample as well (Table III).

When examining measurement invariance by samples and by sexes, model fit statistics supported configural, and metric invariance across studies and configural, metric, and scalar invariance across sex (Table IV). This indicates that in both the THT and GuwM samples, each of the 10 CESDR-10 items contributed in the same way to the 1-factor structure. It also suggests that respondents in these two samples may have systematically different average

scores across items, depending on which sample they were a part of. Across sex, it appears that full invariance of the CESDR-10 was supported, indicating that the scale performs equally well in males and females.

Item Response and Differential Item Functioning Analysis

Table V displays the item discrimination parameters (factor loadings) for all CESDR-10 items by sample and sex separately. Overall, notable differences were not observed between discrimination parameters on any of the items. Results from the MIMIC models indicated statistically significant associations between sample membership and sex on certain items, but the magnitude of the effect was quite small. The largest effect was observed between sex and item 3 ("I felt sad"), for which girls has 0.2 times the odds of endorsing the item compared to boys. These results suggest that the items on the CESDR-10 showed no clinically significant DIF based on sample membership or sex.

Discussion

Data from two national studies of US adolescents suggest generally strong psychometric properties for the CESDR-10. CFA results indicated good model fit of a 1-factor model in the THT and combined samples, and marginal model fit in the GuwM sample. As indicated by examination of measurement correlation and measurement invariance, construct validity was mostly supported in Sample 2 and partially supported in Sample 1. Results indicated that the CESDR-10 is invariant across sex. The IRT and DIF analyses showed no major differences in item discrimination parameters and no clinically significant DIF by sample membership or sex.

Based on an adapted method for scoring (Eaton, Ybarra et al. 2012), 6.4% of the adolescents in the THT sample and 4.8% of adolescents in the GuwM sample met criteria for possible, probable or actual current major depressive episode. Overall, 19.9% of the THT sample and 16.5% of the GuwM sample showed current clinical and subclinical depressive symptoms. In both study samples, prevalence rates for female adolescents were higher for possible, probable or actual current major depressive episode. These estimates of current depressive and subclinical symptomatology, and rates across genders, are consistent with the epidemiological literature on 1-year prevalence rates of adolescent depression (Thapar, Collishaw et al. 2012).

Of particular note, the CESDR-10 includes the item "I felt irritable," which is not included in other revised versions of the CES-D. The factor analysis results suggest a high factor loading of this item on the underlying factor. This represents the strength of the CESDR-10 compared to other 10-item versions of the CES-D because it parallels current criteria for depression in adolescents. Sex and age differences further suggest that the scale could be used for adult populations, with the exclusion of the irritability item.

The measure of social support in the THT study revealed intriguing differences between online and offline support: social support from in-person friends was weakly and negatively correlated with higher scores on the CESR-10, whereas social support from online friends was more strongly and positively correlated with total CESDR-10 scores. This is consistent

with previous research that noted youth with depressive symptomatology were more likely than their non-depressed peers to use the Internet to communicate with others – including those known in-person as well as those known only online (Ybarra, Alexander et al. 2005). It may be that youth who are depressed use the Internet as a way to replace social support that they would otherwise have in person. The Internet may provide opportunities to continue to engage with people at a pace and intensity that is more approachable for depressed youth than in person interactions.

Limitations

This psychometric evaluation had several limitations. First, there is no way to assess criterion related validity. Further studies of the psychometric properties of the CESDR-10 could include the assessment of the scale compared to a psychiatric diagnosis. Another potential limitation of this study was the web-based format of administration. Study setting and administration method may influence participants' answers. To address this limitation the CESDR-10 should be administered in other settings as well, to more thoroughly establish its generalizability.

Conclusions

In clinical and non-clinical settings alike, time pressures make the availability of brief but valid screening measures critical. The CESDR-10 is a psychometrically sound self-report depression scale that should be considered for use in adult as well as adolescent populations in the United States. The use of two different national samples enhances the generalizability of these findings and support the use of this scale in future studies.

Acknowledgements

None

Role of the funding source: The first author is supported by NIMH grant 5T32MH014592-37

References

Allgaier AK, Frühe B, Pietsch K, Saravo B, Baethmann M, Schulte-körne G. Is the Children's Depression Inventory Short version a valid screening tool in pediatric care? A comparison to its full-length version. Journal of psychosomatic research. 2012

American Psychiatric Association. Diagnostic and statistical manual of mental disorders: DSM-IV-TR. American Psychiatric Publishing, Inc; 2000.

Andresen EM, Malmgren JA, Carter WB, Patrick DL. Screening for depression in well older adults: Evaluation of a short form of the CES-D. American Journal of Preventive Medicine. 1994

Bachman, JG.; Johnston, LD.; O'malley, PM. Monitoring the Future: Questionnaire responses from the nation's high school seniors, 2000. Ann Arbor, MI: MI: Institute for Social Research; 2001.

Beck AT, Ward CH, Mendelson M, Mock J, Erbaugh J. An inventory for measuring depression. Archives of General Psychiatry. 1961; 4(6):561. [PubMed: 13688369]

Brener ND, Collins JL, Kann L, Warren CW, Williams BI. Reliability of the youth risk behavior survey questionnaire. American Journal of Epidemiology. 1995; 141(6):575–580. [PubMed: 7900725]

Burnam MA, Wells KB, Leake B, Landsverk J. Development of a brief screening instrument for detecting depressive disorders. Medical care. 1988:775–789. [PubMed: 3398606]

Cartierre N, Coulon N, Demerval R. Confirmatory factor analysis of the short French version of the Center for Epidemiological Studies of Depression Scale (CES-D10) in adolescents. L'Encephale. 2011; 37(4):273–277.

Cheng ST, Chan A, Fung HH. Factorial structure of a short version of the Center for Epidemiologic Studies Depression Scale. International journal of geriatric psychiatry. 2006; 21(4):333–336. [PubMed: 16570325]

Cole JC, Rabin AS, Smith TL, Kaufman AS. Development and validation of a Rasch-derived CES-D short form. Psychological assessment. 2004; 16(4):360. [PubMed: 15584795]

Dahlberg, LL.; Toal, SB.; Swahn, M.; Behrens, CB. Measuring violence-related attitudes, behaviors, and influences among youths: A compendium of assessment tools. Atlanta, GA: Centers for Disease Control and Prevention, National Center for Injury Prevention and Control; 2005.

Derogatis LR, Lipman RS, Rickels K, Uhlenhuth EH, Covi L. The Hopkins Symptom Checklist (HSCL): A self-report symptom inventory. Behavioral science. 1974; 19(1):1–15. [PubMed: 4808738]

Eaton WW, Smith C, Ybarra M, Muntaner C, Tien A. Center for Epidemiologic Studies Depression Scale: review and revision (CESD and CESD-R). 2004

Eaton WW, Ybarra M, Schwab J. The CESD-R is available on the web. Psychiatry research. 2012; 196(1):161. [PubMed: 22349650]

Ellis AA, Nixon RDV, Williamson P. The effects of social support and negative appraisals on acute stress symptoms and depression in children and adolescents. British Journal of Clinical Psychology. 2009; 48(4):347–361. [PubMed: 19254448]

Finch H. The MIMIC model as a method for detecting DIF: Comparison with Mantel-Haenszel, SIBTEST, and the IRT likelihood ratio. Applied Psychological Measurement. 2005; 29(4):278–295.

Finkelhor D, Mitchell KJ, Wolak J. Online Victimization: A Report on the Nation's Youth. 2000

Fletcher JM. Adolescent depression and educational attainment: results using sibling fixed effects. Health Economics. 2010; 19(7):855–871. [PubMed: 19582699]

Gallo JJ, Anthony JC, Muthen BO. Age differences in the symptoms of depression: A latent trait analysis. Journal of gerontology. 1994; 49(6):P251. [PubMed: 7963280]

Garber J, Quiggle NL, Panak W, Dodge KA. Aggression and depression in children: Comorbidity, specificity, and social cognitive processing. Internalizing and externalizing expressions of dysfunction. 1991; 2:225–264.

Goodman R. The Strengths and Difficulties Questionnaire: a research note. Journal of child psychology and psychiatry. 1997; 38(5):581–586. [PubMed: 9255702]

Hamilton M. A rating scale for depression. Journal of neurology, neurosurgery, and psychiatry. 1960; 23(1):56.

Irwin M, Artin KH, Oxman MN. Screening for depression in the older adult: criterion validity of the 10-item Center for Epidemiological Studies Depression Scale (CES-D). Archives of Internal Medicine. 1999; 159(15):1701. [PubMed: 10448771]

Jane costello E, Erkanli A, Angold A. Is there an epidemic of child or adolescent depression? Journal of Child Psychology and Psychiatry. 2006; 47(12):1263–1271. [PubMed: 17176381]

Keenan-miller D, Hammen CL, Brennan PA. Health outcomes related to early adolescent depression. Journal of Adolescent Health. 2007; 41(3):256–262. [PubMed: 17707295]

Kilpatrick DG, Ruggiero KJ, Acierno R, Saunders BE, Resnick HS, Best CL. Violence and risk of PTSD, major depression, substance abuse/dependence, and comorbidity: results from the National Survey of Adolescents. Journal of consulting and clinical psychology. 2003; 71(4):692. [PubMed: 12924674]

Kohout FJ, Berkman LF, Evans DA, Cornoni-huntley J. Two shorter forms of the CES-D depression symptoms index. Journal of aging and health. 1993; 5(2):179–193. [PubMed: 10125443]

Kovacs M. Children's depression inventory (CDI), 2005, MHS. 2005

Kroenke K, Spitzer RL, Williams JB. The Patient Health Questionnaire-2: validity of a two-item depression screener. Medical care. 2003; 41(11):1284–1292. [PubMed: 14583691]

Kroenke K, Spitzer RL, Williams JB. The PHQ-9. Journal of general internal medicine. 2001; 16(9): 606–613. [PubMed: 11556941]

Lin HC, Tang TC, Yen JY, Ko CH, Huang CF, Liu SC, Yen CF. Depression and its association with self-esteem, family, peer and school factors in a population of 9586 adolescents in southern Taiwan. Psychiatry and clinical neurosciences. 2008; 62(4):412–420. [PubMed: 18778438]

Messer SC, Angold A, Costello EJ, Loeber R, Van kammen W, Stouthamer-loeber M. Development of a short questionnaire for use in epidemiological studies of depression in children and adolescents: Factor composition and structure across development. International Journal of Methods in Psychiatric Research. 1995; 5:251–262.

Milfont TL, Merry S, Robinson E, Denny S, Crengle S, Ameratunga S. Evaluating the short form of the Reynolds Adolescent Depression Scale in New Zealand adolescents. Australian and New Zealand Journal of Psychiatry. 2008; 42(11):950–954. [PubMed: 18941959]

Murray CJ, Jamison DT, Lopez AD, Ezzati M, Mathers CD. Global burden of disease and risk factors. 2006

Muthén L, Muthén B. Mplus. The comprehensive modelling program for applied researchers: user's guide. 2012; 5

Nuevo R, Dunn G, Dowrick C, Vázquez-barquero JL, Casey P, Dalgard OS, Lehtinen V, Ayuso-mateos JL. Cross-cultural equivalence of the Beck Depression Inventory: A five-country analysis from the ODIN study. Journal of affective disorders. 2009; 114(1):156–162. [PubMed: 18684511]

Orth U, Robins RW, Roberts BW. Low self-esteem prospectively predicts depression in adolescence and young adulthood. Journal of personality and social psychology. 2008; 95(3):695. [PubMed: 18729703]

Park LQ, Gross AL, Mclaren DG, Pa J, Johnson JK, Mitchell M, Manly JJ. Confirmatory factor analysis of the ADNI neuropsychological battery. Brain imaging and behavior. 2012; 6(4):528–539. [PubMed: 22777078]

Pickles A, Aglan A, Collishaw S, Messer J, Rutter M, Maughan B. Predictors of suicidality across the life span: the Isle of Wight study. Psychological medicine. 2010; 40(9):1453–1466. [PubMed: 19939326]

Price SD, Salekin RT, Klinger MR, Barker ED. Psychopathy and Depression as Predictors of Psychosocial Difficulties in a Sample of Court Evaluated Adolescents. 2012

Radloff LS. The use of the Center for Epidemiologic Studies Depression Scale in adolescents and young adults. Journal of Youth and Adolescence. 1991; 20(2):149–166. [PubMed: 24265004]

Radloff LS. The CES-D scale a self-report depression scale for research in the general population. Applied psychological measurement. 1977; 1(3):385–401.

Reynolds, WM. Reynolds adolescent depression scale. Wiley Online Library; 1986.

Rosenberg, M. Society and the adolescent self-image (rev. Wesleyan University Press; 1989.

Samejima F. Graded response model. Handbook of modern item response theory. 1997:85–100.

Santor DA, Coyne JC. Shortening the CES–D to improve its ability to detect cases of depression. Psychological assessment. 1997; 9(3):233.

StataCorp. Stata Statistical Software: Release 11. College Station, TX: StataCorp LP; 2009.

Steinmetz H, Schmidt P, Tina-booh A, Wieczorek S, Schwartz SH. Testing measurement invariance using multigroup CFA: Differences between educational groups in human values measurement. Quality & Quantity. 2009; 43(4):599–616.

Stringaris A, Zavos H, Leibenluft E, Maughan B, Eley T. Adolescent irritability: phenotypic associations and genetic links with depressed mood. The American Journal of Psychiatry. 2012; 169(1):47. [PubMed: 22193524]

Thapar A, Collishaw S, Pine DS, Thapar AK. Depression in adolescence. The Lancet. 2012; 379(9820):1056–1067.

Udry, JR. The National Longitudinal Study of Adolescent Health (AddHealth) [Wave I and Wave II]. Chapel Hill, NC: Carolina Population Center, University of North Carolina at Chapel Hill; 1996.

Van dam NT, Earleywine M. Validation of the Center for Epidemiologic Studies Depression Scale—Revised (CESD-R): Pragmatic depression assessment in the general population. Psychiatry research. 2011; 186(1):128–132. [PubMed: 20843557]

Ybarra ML, Alexander C, Mitchell KJ. Depressive symptomatology, youth Internet use, and Online interactions: a national survey. Journal of Adolescent Health. 2005; 36(1):9–18. [PubMed: 15661591]

Yu, CY.; Muthen, B. Evaluation of model fit indices for latent variable models with categorical and continuous outcomes. annual meeting of the American Educational Research Association; New Orleans, LA. 2002. 2002

Zimet GD, Dahlem NW, Zimet SG, Farley GK. The multidimensional scale of perceived social support. Journal of personality assessment. 1988; 52(1):30–41.

Table I

Demographic characteristics

	Sample 1: TEEN HEALTH AND TECHNOLOGY (THT)	**Sample 2: GROWING UP WITH MEDIA (GuwM)**
Number of subjects	3777	1150
Age in years; mean (SD)	15.7 (1.65)	14.5 (1.76)
Biological Sex		
Male (%)	1639 (43.4)	582 (50.6)
Female (%)	2138 (56.6)	568 (49.4)
Race		
White (%)	2841 (75.2)	854 (74.3)
Black/African-American (%)	316 (8.4)	150 (13.0)
Mixed (%)	238 (6.3)	83 (7.2)
All other (%)	382 (10.1)	63 (5.5)
Current grade		
5^{th}–8^{th} grade	903 (23.9)	405 (15.5)
9^{th}–12^{th} grade	2721 (72.0)	721 (62.7)
Other	153 (4.1)	24 (2.1)

Table II

Standardized CFA Factor loadings of CESDR-10 items

	Combined Sample	THT (λ)	GuwM (λ)
My appetite was poor	0.75	0.77	0.68
My sleep was restless	0.75	0.76	0.72
I felt sad	0.83	0.84	0.79
I felt like a bad person	0.83	0.83	0.79
I lost interest in my usual activities	0.85	0.87	0.81
I felt like I was moving too slowly	0.80	0.81	0.77
I wished I were dead	0.86	0.87	0.82
I was tired all the time	0.78	0.79	0.77
I could not focus on the important things	0.82	0.82	0.82
I felt irritable	0.79	0.79	0.80
Internal consistency reliability		α	α
		0.91	0.90

Table III

Average depression scores by age and sex

	THT		GuwM	
	Males Mean (S.D.) *range*	**Females Mean (S.D.) *range***	**Males Mean (S.D.) *range***	**Females Mean (S.D.) *range***
11–12 year olds [a]	--	--	14.2 (5.3) *10–37*	15.1(6.2) *10–45*
13–14 year olds [b]	5.2 (7.4) *0–40*	9.4 (10.6) *0–40*	13.6 (4.5) *10–38*	16.8 (8.6) *10–50*
15–17 year olds [c]	7.9 (8.8) *0–40*	10.4 (9.7) *0–40*	14.5 (6.0) *10–45*	15.2 (6.5) *10–50*
18 year olds [d]	7.7 (9.0) *0–40*	9.7 (9.4) *0 – 40*	14.7 (7.3) *10–50*	16.5 (7.2) *10–50*

[a] THT: *n*=0 males, *n*=0 females; GuwM: *n*=95 males, *n*=99 females

[b] THT: *n*=585 males; *n*=785 females; GuwM: *n*=201 males, *n*=191 females

[c] THT: *n*=856 males, *n*=1096 females; GuwM: *n*=192 males; *n*=172 females

[d] THT: *n*=954 males *n*=1404 females; GuwM: *n*=94 males; *n*=106 females

Table IV

Measurement invariance by study sample and sex

Model	Reference Model	Difference testing					RMSEA	CFI	TLI	WRMR
		χ2	*df*	*P*	Δ χ2	*P*Δ				
Study Sample										
1. Configural		718.2	70	0.00	--	--	0.061	0.99	0.98	2.024
2. Metric	1	734.7	79	0.00	16.5	0.00	0.058	0.99	0.99	2.128
3. Scalar	2	614.6	108	0.00	120.1	0.05	0.044	0.99	0.99	2.218
Sex										
1. Configural		700.6	70	0.00	--	--	0.060	0.99	0.98	1.975
2. Metric	1	817.9	79	0.00	117.3	0.00	0.062	0.99	0.98	2.197
3. Scalar	2	733.0	108	0.00	84.9	0.00	0.048	0.99	0.99	2.354

Table V

IRT analysis by study and by sex

Items	GuwM	THT	Males	Females
	λ *(S.E.)*	λ *(S.E.)*	λ *(S.E.)*	λ *(S.E.)*
My appetite was poor	0.68 (0.03)	0.77 (0.01)	0.76 (0.01)	0.74 (0.01)
My sleep was restless	0.72 (0.02)	0.76 (0.01)	0.75 (0.01)	0.74 (0.01)
I felt sad	0.79 (0.02)	0.84 (0.01)	0.83 (0.01)	0.82 (0.01)
I felt like a bad person	0.79 (0.02)	0.83 (0.01)	0.82 (0.01)	0.82 (0.01)
I lost interest in my usual activities	0.81 (0.02)	0.87 (0.01)	0.84 (0.01)	0.86 (0.01)
I felt like I was moving too slowly	0.77 (0.02)	0.81 (0.01)	0.79 (0.01)	0.81 (0.01)
I wished I were dead	0.82 (0.03)	0.87 (0.01)	0.86 (0.01)	0.86 (0.01)
I was tired all the time	0.77 (0.02	0.89 (0.01)	0.77 (0.01)	0.79 (0.01)
I could not focus on the important things	0.82 (0.02)	0.82 (0.01)	0.80 (0.01)	0.83 (0.01)
I felt irritable	0.80 (0.01)	0.79 (0.01)	0.79 (0.01)	0.78 (0.01)