Published in final edited form as:

J Psychopathol Behav Assess. 2019 ; 41(4): 692–698. doi:10.1007/s10862-019-09746-x.

A Factor Analysis and Test of Longitudinal Measurement Invariance of the Children's Depression Inventory (CDI) Across Adolescence

Allison Stumper[1], **Thomas M. Olino**[1], **Lyn Y. Abramson**[2], **Lauren B. Alloy**[1]

[1]Department of Psychology, Temple University

[2]Department of Psychology, University of Wisconsin-Madison

Abstract

Depression increases dramatically during adolescence. This finding has been demonstrated using multiple measures, including the Children's Depression Inventory (CDI). The CDI is one of the most commonly used measures to assess depression in youth. However, there is little agreement on its factor structure, and it is possible that its factor structure changes over time. Yet, no study to date has investigated whether this structure is longitudinally invariant from early- to mid-adolescence. The present study examined the factor structure of the CDI in a sample of 227 adolescents aged approximately 13 at baseline and 16 at follow-up. The analyses revealed that a one-factor structure was a good fit to the data at each assessment. Moreover, tests of measurement invariance supported configural, metric, and scalar invariance across time. These findings suggest that changes in depressive symptoms during adolescence are due to true developmental changes, rather than changes in measurement properties.

Keywords

Factor Analysis; Longitudinal Measurement Invariance; Children's Depression Inventory; Adolescence

One of the most commonly used measures to assess depressive symptoms in youth is the Children's Depression Inventory (CDI; Kovacs, 1985). The CDI was designed as a measure of depression symptom severity that assesses multiple dimensions of depression. The CDI has demonstrated good reliability, discriminant validity, distinguishing children with general distress from other healthy children, and convergent validity based on associations with psychiatrists' ratings of children's symptoms and a self-report measure of attributional style (Saylor, et al., 1984; Masip et al., 2010). The CDI is validated for use in youth from ages 7 to 17 (Sitarenios & Kovacs, 1999). However, these norms presuppose that the structure and

Corresponding author: Lauren B. Alloy, PhD; lalloy@temple.edu; 1701 N. 13th Street, Dept of Psychology, Philadelphia, PA 19122.

Conflict of Interest: All authors declare that they have no conflicts of interest.

Ethical Approval: All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed Consent: Informed consent was collected from all individual participants included in this study.

item properties of the measure are consistent over time. These assumptions remain unexamined, despite their important implications. If the measurement structure and/or functioning of items changes over time, it is not possible to discern whether observed changes in depression are due to “true” changes in depression, or to changes in the measurement of the construct over time.

The CDI is an extrapolation of the Beck Depression Inventory, modified with language for youth (Kovacs, 1992). It was designed to assess five factors (anhedonia, negative mood, negative self-esteem, ineffectiveness, and interpersonal problems) contributing to a total depression score representing symptom severity (Kovacs, 1985). Indeed, these factors represent many of the domains necessary for a diagnosis of Major Depressive Disorder, including depressed or irritable mood, markedly diminished pleasure in all or most activities (including socializing), cognitive impairment (difficulty with decision making, concentration or slowed thinking), and feelings of worthlessness (American Psychological Association, 2013). However, studies using exploratory factor analysis (EFA) have failed to find consistent factor structures for this measure. For example, one study found that a two-factor model, reflecting enjoyment/interpersonal relations and affective behavioral aspects of depression, fit the data (Spence & Milne, 1987), and another found that both two- and three-factor models fit the data well (Carey et al., 1987). In this study, the two-factor model had factors of depressive affect and oppositional behavior, and the three-factor model had the same two factors plus a personal adjustment factor. Hodges et al. (1983) conducted an EFA and concluded that a four-factor model reflecting cognitive, motivational, social integration, and somatic components of depression fit the data best in a clinical sample of children, and a two-factor model comprised of one “general” factor and a second non-compliant behavior factor fit the data in a non-clinic school sample. A fourth study found support for eight and seven factors in samples of healthy and clinic-referred children, respectively (Saylor et al., 1984). A fifth study implementing only a confirmatory factor analysis (CFA) found evidence for the original five-factor structure in both European-American and African-American children (Steele, Little, Ilardi, Forehand, Brody, & Hunter, 2006). Finally, one study obtained a unidimensional factor structure of the CDI in a sample of American Indian and Native Alaskan youth in grades 4–12 (Scott, Clapp, Mileviciute, & Mousseau, 2016).

The differences in factor structures are especially apparent when comparing samples of children and adolescents. For instance, after conducting an EFA and CFA in a large community sample of youth from 8 to 17 years, Craighead et al. (1998) obtained five factors for children (externalizing, dysphoria, self-deprication, school problems, social problems) and the same five factors plus an additional sixth factor (biological dysregulation) in adolescents. Subsequent CFAs found support for the same 5- and 6-factor structures, as well as two higher order factors (internalizing and externalizing).

Weiss et al. (1991) found evidence for five similar factors for children and adolescents using EFA. A subsequent CFA supported this five-factor model and also found evidence for a single higher order “depression” factor; however, the factor loadings from the depression factor to observed indicators differed across children and adolescents. Importantly, this study reported that a preliminary scree test indicated the best fitting model for both children and adolescents is a single factor.

Some studies, however, have concluded that the factor structure of the CDI is the same for children and adolescents. Garcia, Aluja, and Barrio (2008) conducted both an EFA and CFA in a large sample of Spanish children, guided by findings from previous studies. They found that a model with five first-order factors and one higher order factor fit the data well for both children (aged 7–12) and adolescents (aged 13–16).

In sum, there is limited agreement on the factor structure of the CDI in the literature, and differences in factor structure may exist across developmental stages. To our knowledge, only two studies to date have directly compared the factor structure of the CDI across development (Weiss et al., 1991; Garcia et al., 2008). However, the samples used were broad and heterogeneous in age. Both studies defined youth aged 7- or 8- to 12-years old as children and youth aged 13–16 years old as adolescents. There are substantial developmental differences between a 7- and 12-year old, as well as between a 13- and 16-year old. More work is needed to understand whether there are differences in the structure of the CDI across each stage of development, especially from early to later adolescence.

The ages of 13 and 15–16 are particularly relevant to the study of depression. Many studies have found that increases in depression begin around age 13, and the preponderance of depression in females is evident by age 15–16 (Hankin & Abramson, 2001; Hankin, Abramson, Moffitt, Silva, & McGee, 1998). These findings have been replicated using measures of depressive symptoms, like the Symptom Checklist-90 (SCL-90), the Children's Depression Inventory (CDI), the Kandel Depression Scale, the Child Behavior Checklist (CBCL) and the Youth Self-Report (YSR) (Ge, Conger, & Elder, 2001; Cole, Martin, Peeke, Seroczynski & Fier, 1999; Petersen et al., 1991; Wichstrom, 1999; Compas et al., 1997) and diagnoses, like the Diagnostic Interview Schedule (DIS) and the Kiddie Schedule for Affective Disorders and Schizophrenia (K-SADS) (Hankin et al., 1998; Lewinsohn, 1993; Cohen, et al., 1993). However, little work has investigated the stability of the measurement structure across adolescence in many of the most commonly used instruments for assessing depressive symptoms and diagnoses, like the CDI. There are significant biological, cognitive, emotional and social differences in youth of these ages (Steinberg, 2010; Yurelun-Todd, 2007; Brook-Gunn & Warren, 1989), all of which could influence their report of depressive symptoms and structure of responses on the CDI across time. Differences in the structure of depression across development would challenge our understanding of mean level changes in symptoms across time. Thus, in order to have confidence that the developmental changes observed are attributable to actual changes in individuals across time and not measurement error, formal tests for longitudinal measurement invariance in depressive measures are needed. The current study first conducted an EFA to determine the best fitting factor structure for the CDI in a sample diverse in gender, race, and socioeconomic status; we then tested for measurement invariance in the CDI between individuals aged 12–13 and aged 15–16 in this sample.

Methods

Participants

The study sample was drawn from the Adolescent Cognition and Depression (ACE) Project, an ongoing prospective, longitudinal study of risk factors for the development of depression

during adolescence. Recruitment was conducted through either school mailings and follow-up phone calls inviting the mother/primary caregiver and her child to participate or through advertisements in Philadelphia-area newspapers. Eligibility required that the adolescent was 12 or 13 years old, identified as Caucasian/White, African American/Black, or Biracial, and the mother/primary caregiver also agreed to participate. Individuals who identified as other ethnicities were excluded because an aim of the study was to examine differences in cognitive vulnerabilities to depression between Black and White adolescents. Exclusion criteria were the absence of a mother/primary caregiver, either the child or the mother's English was insufficient to complete study assessments, and either the adolescent or mother was psychotic, intellectually disabled, or had a pervasive developmental disorder or learning disability.

The current study sample consisted of 227 early adolescents (M_{age}=12.88, SD=.65 at baseline) with complete data at baseline (T1) and follow-up (T2). At T2, participants were approximately 16 years old (M_{age}=16.03, SD=.71). In the current sample, 51.5% of the adolescents identified as female, 55.9% identified as African-American, and 42.7% were eligible for subsidized school lunch, which takes into account the number of dependents supported on the family's income and was a proxy for socioeconomic status (SES). SES and race were not significantly correlated with depressive symptoms at T1 or T2 (SES: r=.09, p=.17, r=.12, p=.08, respectively; race: r=−.02, p=.77, r=.07, p=.31, respectively). Females reported significantly higher depressive symptoms at both T1 and T2 (r=−.13, p=.04, r=.17, p=.01, respectively).

The current sample did not differ from the larger study sample on sex (t(639)=1.56, p=.12), race (t(639)=.99, p=.32), or CDI at T1 (t(631)=1.18, p=.24), but they were less likely to be eligible for subsidized school lunch (a proxy for low SES; t(610)=.21, p=.04).

Procedures

Participants in the present study were assessed at two timepoints. The goal of the ongoing ACE Project is to interview participants every six months. At each follow-up, participants complete a battery of self-report questionnaires, interviews, and behavioral tasks. The CDI is given at every follow-up. Data for the current project were drawn from participants' baseline assessment (T1) and a routine assessment three years later (M=3.15 years, SD=.27).

Measures

Depressive Symptoms.—The Children's Depression Inventory (CDI; Kovacs, 1985) was used to assess depressive symptoms. This self-report measure consists of 27 items, each scored on a scale of 0 to 2. Participants choose the statement that best describes how they have been feeling for the past two weeks. An example item is "I am sad once in a while/I am sad many times/I am sad all the time." For some items on the CDI, a higher score indicates a more severe symptom, whereas for other items, a lower score indicates a more severe symptom. Prior to running analyses, we reverse scored items where lower scores indicated more severe symptoms; so for all items, a higher score indicated more severe symptoms. This allowed us to interpret a positive correlation with a factor to mean that the item was significantly positively associated with that factor without referring to the content of that

item. Thus, all items were expected to have positive factor loadings. On some items, there was limited to no endorsement of response options that indicated elevated symptoms. Three items at T1 (2, 3, and 9) and six items at T2 (1, 3, 4, 9, 22 and 27) had the options indicating the most severe depressive symptom endorsed one or zero times. Thus, data were recoded to be dichotomous (0/1, not endorsed/endorsed). The CDI has been demonstrated to have good construct validity (Doerfler, Felner, Rowlison, Raley, & Evans, 1988), to predict depressive disorders, and to differentiate depressive disorders from anxiety disorders and externalizing disorders among youth (Timbremont, Braet, & Dreessen, 2004). Internal consistency for this measure was adequate at T1 and T2 (α=.85 and α=.87, respectively).

Data Analysis

All analyses were completed in Mplus Version 7 (Muthén, L. K., & Muthén, B. O. (1998–2017). *Mplus User's Guide. Seventh Edition*. Los Angeles, CA: Muthén & Muthén.). As previous studies have reported discrepant factor structures for the CDI, we first conducted EFA to identify the structure in this sample and then fit CFA models to test for longitudinal invariance. We evaluated model fit based on the comparative fit index (CFI), Tucker-Lewis index (TLI) and root mean square error of approximation (RMSEA). Good model fit is indicated by a CFI ≥ .95, TLI > .95, and a RMSEA ≤ .05. Acceptable fit is indicated by a CFI > .90, TLI > .90, and a RMSEA < .08. We did not rely on the chi square value as an indicator of fit, as the chi square is known to be sensitive to large sample sizes (Schermelleh-Engel et al., 2003).

To test longitudinal measurement invariance, we implemented a CFA framework. This model specified individual depression factors at ages 12/13 and 15/16 and permitted these factors to be correlated. We followed three steps, each imposing an additional restriction on the measurement model. Model fit information was compared with that of the previous, less restricted model. Invariance was considered to be met if the fit of the more restrictive model was not significantly less than that of the less restricted model (Widaman, Ferrer, & Conger, 2010).

The first step of longitudinal measurement invariance is configural invariance, which imposes the same pattern of factor loadings at each timepoint, but does not make any equality constraints. We tested configural invariance by specifying the CDI factors at each timepoint and permitted correlations among factors across time. The second level of longitudinal measurement invariance is metric invariance, which tests whether factor loadings to each observed indicator are equal across timepoints. Thus, to test metric invariance, we constrained factor loadings to be equal at age 12/13 and age 15/16. The third level of longitudinal measurement invariance is scalar invariance, which tests whether the item thresholds are invariant across time. In other words, it tests whether the likelihood of endorsing each item is equivalent at each timepoint. Support for scalar invariance indicates that there are no inherent differences in the likelihood of endorsing items between each timepoint, and therefore, that changes in depression scores on the CDI reflect a true change in depression levels over time, not just the effect of age.

When testing measurement invariance, we used a guideline of a change in either the CFI, TFI, or RMSEA of ≥ .01 to be indicative of significant decrement in model fit (Chen, 2007).

Per recommendations by Muthen, du Toit, and Spisic (2013), a weighted least square mean and variance (WLSMV) estimation method was used in Mplus (Muthen and Muthen, 2005) because observed variables were binary.

Results

Exploratory Factor Analyses of the CDI at Ages 13 and 16[1]

An EFA using an oblique oblimin rotation was conducted to evaluate the best fitting model for each timepoint. At both age 12/13 and age 15/16, eigenvalues between a 1-factor solution and 2-factor solution had a ratio > 4:1 (T1: 1 factor = 11.843 & 2 factor = 1.993; T2: 1 factor = 9.728 & 2 factor = 2.335), which suggests that this measure captures a unidimensional measure of depression (Slocum-Gori & Zumbo, 2011). In addition, other model fit indices indicated that model fit was good (T1: CFI = 0.94, TLI = 0.94, RMSEA = .03; T2: CFI = 0.94, TLI = 0.94, RMSEA = .04). Model fit for models with 2, 3, and 4-factors, was also good, and in some cases, better than a 1-factor model. However, these factor structures had questionable interpretations. First, the items that significantly loaded onto each factor did not represent a unitary construct. Second, there were multiple cross-loadings for items, which further diminished the interpretations of the scales. Therefore, a one factor model was used to test measurement invariance. A CFA was conducted to test the fit of a 1-factor model for each timepoint. Model fit was good at both ages (Table 1) and factor loadings at each timepoint are presented in Table 1.

Configural Invariance

Given the good fit of the one-factor model for both ages, measurement invariance testing continued. In the configural models, we specified one factor at each time point and permitted correlations among factors across time. The CFI and RMSEA for this model were both good (Table 2), and all factor loadings to their respective latent variables were significant and in the expected direction (Table 1). This indicates that the same factor structure of the CDI fits the data at each timepoint.

Metric Invariance

To test metric invariance, items had loadings to the single factor, and factor loadings for corresponding items were constrained to be equal across time. Overall model fit was good (Table 2). Metric invariance was supported because the changes in RMSEA, CFI, and TLI were < .01 (Table 2). This implies that the factor loadings to items on the CDI do not differ between age 12/13 and age 15/16.

Scalar Invariance

To test scalar invariance, we imposed equality of thresholds across the two timepoints in the model, allowing us to test whether the likelihood of response endorsement can be attributed

[1]We also tested a bifactor rotation that modeled a general factor and up to four specific factors. In each of these models, there was evidence of a strong general factor, with factor loadings being substantial and significant at $p < .05$. However, there was very questionable evidence of the utility and interpretability of the specific factors. Specific factors had few loadings that were significant and there were multiple loadings that were negative. Thus, as the single factor model provided a satisfactory fit to the data, the improved fit of the bifactor model could be due to over fitting (Bonifay, Lane, & Reise, 2017; Markon, 2019).

to the true level of depression of the participant as opposed to age. Overall model fit was good (Table 2). There was a negligible reduction in model fit relative to the metric invariance (i.e., changes in CFI, TLI, and RMSEA were < .01 of those for the metric invariance model). These results suggest that thresholds are equal across the ages.

Discussion

There is a large literature on the increase in depression during adolescence (Hankin et al., 1998). Many studies utilize the CDI to measure depressive symptoms among youth, but the factor structure of this measure has not been highly replicable, and no study has investigated the longitudinal measurement invariance of this measure from early- to mid-adolescence within the same sample of youth. In order to have confidence in longitudinal assessments of adolescent depression and that our findings are due to developmental differences and not changes in measurement properties, it is essential that we evaluate the structure of the measure and test longitudinal measurement invariance.

The current study found support for a one-factor model of depression using the CDI in this sample. Prior work has found a number of discrepant factor solutions, with models ranging from two to eight factors depending on the sample (Craighead et al. 1998; Weiss et al., 1991; Garcia et al., 2008; Spence & Milne, 1987; Carey et al., 1987; Hodges et al., 1987; Saylor et al., 1984). However, several of these studies also found evidence for a single higher order depression factor (Weiss et al., 1991; Garcia et al., 2008). Consistent with the present findings, Weiss et al. (1991) reported that a scree test suggested that a one-factor model may fit the data best, and Scott et al. (2016) found support for a unidimensional model of depression from the CDI. In practice, the CDI typically is treated like a unidimensional measure, with studies reporting total scores and not utilizing subscale scores. Thus, the current study suggests that this is an appropriate way to use this measure.

It is unclear why this is one of the only studies to conclude that a one-factor model fit the data for this measure, but one possibility is that we used different criteria for evaluating how many factors should be retained. This study relied on the ratio between the eigenvalues of the first and second factors, RMSEA, TLI, and CFI. Most previous work has retained any factor with an eigenvalue > 1.0. However, this approach can yield too many factors that lack clinically meaningful constructs (Boyle, 1985). When Scott and colleagues (2016) evaluated the conceptual basis of the two to four factor solutions for the CDI, they concluded that, despite having good model fit, these models represented a response-style artifact or that the factors were poorly defined, and a unidimensional model was more conceptually sound. Indeed, this also could explain the discrepancy in the appropriate number of factors for the CDI in the literature. Alternatively, this finding could be attributed to the fact that the present study changed response options from 3-point scales to binary ones. This led to restriction in range for the analyses and led to a simpler factor structure.

The current study also found support for measurement invariance of the CDI in a sample of youth aged approximately 13 at baseline and 16 at follow-up. Most work has found that age 13 is the time that depression prevalence rates start increasing among girls, and the rate of depression among girls is nearly double that of boys by ages 15–16 (Hankin & Abramson,

2001), making these ages particularly salient to the study of measurement invariance in our measures of depressive symptoms. Our results support the assumption that these findings are due to true developmental changes, and not differences in the measurement of depression across time. In particular, support for scalar invariance allows us to conclude that there are not inherent differences in the likelihood to endorse an item on the CDI between these age groups. Therefore, differences in scores between a 12/13-year old and a 15/16-year old can be attributed to differences in depressive symptoms, not age. Our findings complement existing work that has found that the CDI is invariant across race (Steele et al., 2006) and between children and adolescents (Garcia et al., 2008).

Most factor loadings were comparable from T1 to T2 (Table 1). However, the factor loadings on items 21 (“I never have fun at school/I have fun at school once in a while/I have fun at school many times”) and 22 (“I have plenty of friends/I have some friends but I wish I had more/I do not have any friends”) were quite different between these timepoints. As youth move from childhood to adolescence, they transition from their family as their primary social context to friends. Furthermore, feelings about school may be determined by how much emphasis they place on socializing with friends. It is plausible that most children have not yet made this transition at age 13, but they have by ages 15–16. This could explain why the magnitude of the loadings to these items differs from age 12/13 to age 15/16. However, this difference in the way children value school and social contexts should be considered in the measurement and assessment of depression as youth advance from childhood to adolescence.

This study should be interpreted in light of its limitations. Given that we studied a community sample of adolescents, some options on items were very rarely endorsed. The items in this measure were recoded for the purposes of these analyses to be dichotomous, rather than the original item scaling. Thus, whether these findings generalize to use of the questionnaire in the manner it was written or for more severe samples is unclear. Future work using a larger sample with more variability in symptoms should aim to replicate these findings.

Additionally, these analyses did not consider sex or pubertal development, which both independently and in interaction with each other influence the onset of adolescent depression (Angold, Costello, & Worthman, 1998). The current sample did not have sufficient variability in pubertal development and was not large enough to split by sex, so the question of measurement invariance by sex or stage of pubertal development could not be tested. However, future work should aim to test these questions, as it is essential to the longitudinal measurement of depression.

Finally, relative to the overall sample, those participants included in these analyses were less likely to be eligible for subsidized school lunch. Therefore, the generalizability to low-SES populations should be interpreted cautiously. However, roughly 43% of the current sample was eligible for subsidized school lunch, suggesting that these findings are likely to generalize across different levels of SES.

It also should be noted that the present study discusses and evaluates the CDI, although an updated version of this measure, the 28-item CDI-2 (Kovacs, 2004), is available. While most of the items remain unchanged, some edits were made. Items 6 and 26 were removed from the original CDI. The most symptomatic options of these items read “I believe bad things will happen to me,” and “I never do what I am told.” Items 5, 10, and 25 were edited (e.g., The most symptomatic option on item 10 used to say “I feel like crying everyday” but now says “I feel cranky all the time). Finally, three new items were added to assess hypersomnia, increased appetite, and cognitive difficulties. This new version includes two additional second-order subscales assessing Emotional and Functional Problems and has demonstrated excellent psychometric properties (Bae, 2012). This measure has been utilized and evaluated to a lesser extent than its predecessor, and there is still a need to understand the structure and measurement invariance of the original version. Further, the current study allows us to have confidence in current models of adolescent depression that have been derived from clinical research using the original version of the CDI and in work that may still come from datasets that include the original version. Nevertheless, the field should move towards using more updated instruments, and the present study should be replicated using the CDI-2.

The implications for clinical practice that can be derived from the current study are somewhat limited. There is evidence that the structure of depressive symptoms changes over the course of treatment (Fried et al., 2016), but the sample composition and time frame of the current study (e.g., community sample followed for two years) is quite different from that of a treatment-seeking sample followed over the course of an intervention. Thus, shorter term studies in treatment-seeking youth are needed to evaluate whether changes in CDI scores over the course of treatment reflect true changes in depressive symptoms.

In summary, the current study found evidence for a one-factor model of the CDI, supporting its use as a unidimensional measure in practice. The present findings also offer preliminary evidence for longitudinal measurement invariance of a one-factor model of the CDI from early- to mid-adolescence, adding to the existing literature demonstrating this measure’s sound psychometric properties. Therefore, past work showing an increase in depressive symptoms between ages 13 and 16 is likely reflective of true developmental change. However, more work is needed to test measurement invariance of the CDI by sex, by pubertal development, and in a sample with more variability in symptoms in order to have more confidence in the generalizability of these findings.

Funding:

Preparation of this article was supported by National Institute of Mental Health grants MH79369 and MH101168.

References

American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders (5th ed.). Washington, DC: Author.

Angold A, Costello EJ, & Worthman CM (1998). Puberty and depression: the roles of age, pubertal status and pubertal timing. Psychological medicine, 28(1), 51–61. [PubMed: 9483683]

Bae Y (2012). Test Review: Children’s Depression Inventory 2 (CDI 2). Journal of Psychoeducational Assessment, 30(3), 304–308.

Bonifay W, Lane SP, & Reise SP (2017). Three Concerns With Applying a Bifactor Model as a Structure of Psychopathology. Clinical Psychological Science, 5, 184–186.

Boyle GJ (1985). Self-report measures of depression: Some psychometric considerations. British Journal of Clinical Psychology, 24(1), 45–59. [PubMed: 3971068]

Brooks-Gunn J, & Warren MP (1989). Biological and social contributions to negative affect in young adolescent girls. Child development, 40–55. [PubMed: 2702873]

Carey MP, Faulstich ME, Gresham FM, Ruggiero L, & Enyart P (1987). Children's Depression Inventory: Construct and discriminant validity across clinical and nonreferred (control) populations. Journal of Consulting and Clinical Psychology, 55(5), 755. [PubMed: 3454787]

Chen FF (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. Structural Equation Modeling, 14(3), 464–504.

Cohen P, Cohen J, Kasen S, Velez CN, Hartmark C, Johnson J, … & Streuning EL (1993). An epidemiological study of disorders in late childhood and adolescence—I. Age-and sex-specific prevalence. Journal of child psychology and psychiatry, 34(6), 851–867. [PubMed: 8408371]

Cole DA, Martin JM, Peeke LA, Seroczynski AD, & Fier J (1999). Children's Over-and Underestimation of Academic Competence: A Longitudinal Study of Sex Differences, Depression, and Anxiety. Child development, 70(2), 459–473. [PubMed: 10218266]

Compas BE, Oppedisano G, Connor JK, Gerhardt CA, Hinden BR, Achenbach TM, & Hammen C (1997). Sex differences in depressive symptoms in adolescence: Comparison of national samples of clinically referred and nonreferred youths.

Compian L, Gowen LK, & Hayward C (2004). Peripubertal girls' romantic and platonic involvement with boys: Associations with body image and depression symptoms. Journal of Research on adolescence, 14(1), 23–47.

Craighead WE, Smucker MR, Craighead LW, & Ilardi SS (1998). Factor analysis of the Children's Depression Inventory in a community sample. Psychological Assessment, 10(2), 156.

Doerfler LA, Felner RD, Rowlison RT, Raley PA, & Evans E (1988). Depression in children and adolescents: A comparative analysis of the utility and construct validity of two assessment measures. Journal of consulting and clinical psychology, 56(5), 769. [PubMed: 3192795]

Fried EI, van Borkulo CD, Epskamp S, Schoevers RA, Tuerlinckx F, & Borsboom D (2016). Measuring depression over time… Or not? Lack of unidimensionality and longitudinal measurement invariance in four common rating scales of depression. Psychological Assessment, 28(11), 1354. [PubMed: 26821198]

García LF, Aluja A, & del Barrio V (2008). Testing the hierarchical structure of the Children's Depression Inventory: A multigroup analysis. Assessment, 15(2), 153–164. [PubMed: 18463406]

Ge X, Conger RD, & Elder GH Jr (2001). Pubertal transition, stressful life events, and the emergence of sex differences in adolescent depressive symptoms. Developmental psychology, 37(3), 404. [PubMed: 11370915]

Hankin BL, Abramson LY, Moffitt TE, Silva PA, McGee R, & Angell KE (1998). Development of depression from preadolescence to young adulthood: emerging sex differences in a 10-year longitudinal study. Journal of abnormal psychology, 107(1), 128. [PubMed: 9505045]

Hankin BL, & Abramson LY (2001). Development of sex differences in depression: An elaborated cognitive vulnerability–transactional stress theory. Psychological bulletin, 127(6), 773. [PubMed: 11726071]

Hankin BL, Abramson LY, Moffitt TE, Silva PA, McGee R, & Angell KE (1998). Development of depression from preadolescence to young adulthood: emerging sex differences in a 10-year longitudinal study. Journal of abnormal psychology, 107(1), 128. [PubMed: 9505045]

Hodges KK, Siegel LJ, Mullins L, & Griffin N (1983). Factor analysis of the Children's Depression Inventory. Psychological Reports, 53, 759–763. [PubMed: 6657832]

Kovacs M (1985). The children's depression inventory (CDI). Psychopharmacol bull, 21, 995–998. [PubMed: 4089116]

Kovacs M (1992). CDI. Multi-Health Systems, Inc., North Tonawanda, NY.

Kovacs M (2004). Children's depression inventory (CDI). Toronto, Canada: Multi-Health Systems.

Lewinsohn PM, Hops H, Roberts RE, Seeley JR, & Andrews JA (1993). Adolescent psychopathology: I. Prevalence and incidence of depression and other DSM-III—R disorders in high school students. Journal of abnormal psychology, 102(1), 133. [PubMed: 8436689]

Lubke GH, & Muthén BO (2004). Applying multigroup confirmatory factor models for continuous outcomes to Likert scale data complicates meaningful group comparisons. Structural Equation Modeling, 11(4), 514–534.

Marcotte D, Fortin L, Potvin P, & Papillon M (2002). Sex differences in depressive symptoms during adolescence: Role of sex-typed characteristics, self-esteem, body image, stressful life events, and pubertal status. Journal of Emotional and Behavioral Disorders, 10(1), 29–42.

Markon KE (2019). Bifactor and hierarchical models: Specification, inference, and interpretation. Annual review of clinical psychology, 15, 51–69.

Masip AF, Amador-Campos JA, Gómez-Benito J, & del Barrio Gándara V (2010). Psychometric properties of the Children's Depression Inventory in community and clinical sample. The Spanish journal of psychology, 13(2), 990–999. [PubMed: 20977046]

Muthén LK, & Muthén BO (1998–2017). Mplus User's Guide. Seventh Edition Los Angeles, CA: Muthén & Muthén

Muthén B, Du Toit SHC, & Spisic D (2013). Robust inference using weighted least squares and quadratic estimating equations in latent variable modeling with categorical and continuous outcomes. 1997. Article_075. pdf.

Petersen AC, Compas BE, Brooks-Gunn J, Stemmler M, Ey S, & Grant KE (1993). Depression in adolescence. American Psychologist, 48(2), 155. [PubMed: 8442570]

Petersen AC, Sarigiani PA, & Kennedy RE (1991). Adolescent depression: Why more girls?. Journal of youth and adolescence, 20(2), 247–271. [PubMed: 24265009]

Saylor CF, Finch AJ, Spirito A, & Bennett B (1984). The Children's Depression Inventory: A systematic evaluation of psychometric properties. Journal of consulting and clinical psychology, 52(6), 955. [PubMed: 6520288]

Scott WD, Clapp J, Mileviciute I, & Mousseau A (2016). Children's Depression Inventory: A unidimensional factor structure for American Indian and Alaskan native youth. Psychological assessment, 28(1), 81. [PubMed: 26011483]

Schermelleh-Engel K, Moosbrugger H, Müller H. Evaluating the fit of structural equation models: Test of significance and descriptive goodness-of-fit measures. Methods in Psychology Research. 2003;8(2):23–74.

Sitarenios G, & Kovacs M (1999). Use of the Children's Depression Inventory.

Slocum-Gori SL, & Zumbo BD (2011). Assessing the unidimensionality of psychological scales: Using multiple criteria from factor analysis. Social Indicators Research, 102(3), 443–461.

Steele RG, Little TD, Ilardi SS, Forehand R, Brody GH, & Hunter HL (2006). A confirmatory comparison of the factor structure of the Children's Depression Inventory between European American and African American youth. Journal of Child and Family Studies, 15(6), 773–788.

Spence SH, & Milne J (1987). The Children's Depression Inventory: Norms and factor analysis from an Australian school population. Australian Psychologist, 22(3), 345–351.

Steinberg L (2010). A dual systems model of adolescent risk-taking. Developmental Psychobiology: The Journal of the International Society for Developmental Psychobiology, 52(3), 216–224.

Timbremont B, Braet C, & Dreessen L (2004). Assessing depression in youth: relation between the Children's Depression Inventory and a structured interview. Journal of Clinical Child and Adolescent Psychology, 33(1), 149–157. [PubMed: 15028549]

Twenge JM, & Nolen-Hoeksema S (2002). Age, sex, race, socioeconomic status, and birth cohort difference on the children's depression inventory: A meta-analysis. Journal of abnormal psychology, 111(4), 578. [PubMed: 12428771]

Weiss B, Weisz JR, Politano M, Carey M, Nelson WM, & Finch AJ (1991). Developmental differences in the factor structure of the Children's Depression Inventory. Psychological Assessment: A Journal of Consulting and Clinical Psychology, 3(1), 38.

Wichstrøm L (1999). The emergence of sex difference in depressed mood during adolescence: The role of intensified sex socialization. Developmental psychology, 35(1), 232. [PubMed: 9923478]

Widaman KF, Ferrer E, & Conger RD (2010). Factorial invariance within longitudinal structural equation models: Measuring the same construct across time. Child Development Perspectives, 4(1), 10–18. [PubMed: 20369028]

Yurgelun-Todd D (2007). Emotional and cognitive changes during adolescence. Current opinion in neurobiology, 17(2), 251–257. [PubMed: 17383865]

Table 1.

Factor loadings for a one-factor model of the Children's Depression Inventory at Time 1 and Time 2.

	Time 1		Time 2	
	b	**SE**	**b**	**SE**
CDI1	0.78***	0.08	0.74***	0.07
CDI2	0.58***	0.07	0.60***	0.08
CDI3	0.55***	0.09	0.65***	0.08
CDI4	0.50***	0.08	0.71***	0.06
CDI5	0.68***	0.09	0.67***	0.08
CDI6	0.45***	0.08	0.56***	0.08
CDI7	0.80***	0.07	0.85***	0.06
CDI8	0.59***	0.08	0.77***	0.07
CDI9	0.70***	0.07	0.72***	0.07
CDI10	0.76***	0.06	0.78***	0.06
CDI11	0.73***	0.05	0.74***	0.05
CDI12	0.53***	0.09	0.58***	0.08
CDI13	0.52***	0.08	0.64***	0.07
CDI14	0.56***	0.08	0.66***	0.07
CDI15	0.47***	0.08	0.57***	0.08
CDI16	0.67***	0.07	0.61***	0.07
CDI17	0.74***	0.06	0.64***	0.08
CDI18	0.30**	0.09	0.48***	0.09
CDI19	0.34***	0.09	0.52***	0.08
CDI20	0.83***	0.05	0.81***	0.05
CDI21	0.24***	0.09	0.61***	0.07
CDI22	0.30***	0.10	0.64***	0.08
CDI23	0.46***	0.09	0.63***	0.07
CDI24	0.59***	0.08	0.61***	0.07
CDI25	0.57***	0.09	0.81***	0.07
CDI26	0.56***	0.08	0.47***	0.08
CDI27	0.74***	0.09	0.73***	0.09

Note:

*** $p<.001$,

** $p<.01$,

* $p<.05$.

Table 2.

Model fit indices for baseline models and measurement invariance models.

Model	χ^2	df	p_{χ^2}	CFI	RMSEA	90% CI
T1, only	396.089	324	0.0038	0.941	0.031	.019–.041
T2, only	433.183	324	<.0001	0.943	0.039	.028–.048
Configural	1557.448	1376	0.0004	0.922	0.024	.017–.030
Metric	1592.343	1402	0.0007	0.924	0.024	.016–.030
Scalar	1630.556	1428	0.0003	0.918	0.024	.017–.030

Note: χ^2= chi square; df = degrees of freedom; p_{χ^2}= significance of the chi square statistic; CFI = Comparative Fit Index; RMSEA = Root Mean Square Error of Approximation; 90% CI = 90% confidence interval of the RMSEA.