Children's Depression Inventory: Testing Measurement Invariance for the Hierarchical Factor Model Across Children and Adolescents in a Clinic-Referred Sample

Rapson Gomez[1] and Alasdair Vance[2]

1   Federation University, Australia
2   Royal Children's Hospital, University of Melbourne, Australia

Corresponding author:
Rapson Gomez

✉ rapson.gomez@federation.edu.au

School of Health Sciences and Psychology, Federation University Australia, University Drive, Mt Helen, P.O. Box 663, Ballarat, 3353, Victoria, Australia.

Tel: 0353276087

Abstract

The Children's Depression Inventory is a self-report scale for screening depressive disorders in children and adolescents. The original model proposed by Kovacs has a hierarchical factor structure: Five first-order and a single second-order factors. This study used confirmatory factor analysis (CFA) to examine support for this model. It also examined measurement invariance of this model across self-ratings provided by clinic-referred children (N=459) and adolescents (N=343), and the differences in the first- and second-order latent factor mean scores across these groups. The findings supported the hierarchical factor structure, and also full measurement invariance for this model across the groups compared. Also all latent mean scores were higher in the adolescent group. These findings indicate support for the original CDI model proposed by Kovacs, and also that the ratings provided by clinic-referred children and adolescents can be compared as they are not confounded by different measurement properties. Also, depression is higher among adolescents than children.

Keywords: Children's depression inventory; Children and adolescent ratings; Measurement invariance

Introduction

The Children's Depression Inventory [1] is a 27-item self-report scale, used extensively for assessing child and adolescent depression internationally [2]. The model proposed for the CDI by Kovacs [1], based on exploratory factor analysis (EFA) of ratings provided by children and adolescents from the general community, has a hierarchical structure, with five first-order factors (called Negative Mood, Interpersonal Problems, Ineffectiveness, Anhedonia, and Negative Self-Esteem), and a single second-order factor (called General Depression) **(Figure 1)**. The same factor structure and scoring procedure were proposed for children and adolescents. This assumes that there is measurement invariance across the ratings of the CDI from children and adolescents. The current study used confirmatory factor analysis (CFA) to examine support for Kovacs's hierarchical model for ratings provided by clinic-referred children and adolescents, and also measurement invariance for ratings across these groups. It also tested the differences between children and adolescents for the first- and second-order latent mean scores in this model.

Consistent with the model proposed by Kovacs, EFA studies involving both community and clinic samples [3,4] and CFA studies involving community samples [3,5] have supported a single higher order factor. In contrast, these and other studies [6-8], including the CFA study by Garcia et al. found no support for the first-order Kovacs model. Steele, Little, Ilardi, Forehand, Brody, and Hunter [9] and Logan et al., [5] however found support for this model. Overall, existing support for this model from CFA studies, albeit limited, is mixed. Thus there is a need for more evaluation of the original factor structure proposed by Kovacs. The CDI items responses are order-categorical in nature. Thus it would be useful that future CFA studies use extractions procedures appropriate for such data set, such as the mean and variance-adjusted weighted least squares (WLSMV). The WLSMV is a robust estimator, recommended for CFA with ordered-categorical scores [10].

Most of the existing data suggest different CDI factor models for children and adolescents [3,4,6]. As this level of equivalence, referred to as configural invariance, is a prerequisite for
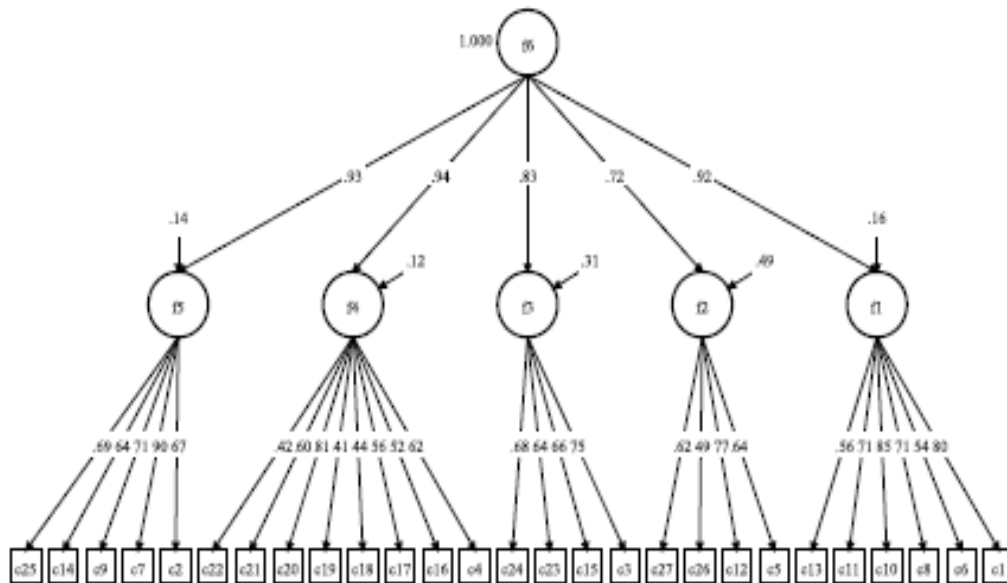
Figure 1 Path diagram with completely standardized estimates of the hierarchical factor model of the CDI for all participants together.

Note: For clarity error variances are not shown. These paths were not correlated.

Item error variance=(1–item factor loading²).

F1: Negative Mood; F2: Interpersonal; F3: Ineffectiveness; F4: Anhedonia; F5: Negative Self-Esteem; F6: General Depression.

measurement invariance, it could mean that CDI items ratings provided by these groups will lack measurement invariance. For a rating scale reflecting a hierarchical factor model, such as the CDI, measurement invariance for the first-order factor model deals with whether the items in the rating scale have the same scale properties when completed by individuals from different groups [11], such as children and adolescents. Measurement invariance for the second-order factor model deals with whether there is group equivalence for their ratings in terms of the relationships of the lower and higher order factors. If there is weak or no support for invariance for the first-order factor and second-order factor models, then it follows that the individuals from the different groups examined cannot be justifiably compared on the raw scores of the first-order factors and second-order factor(s) as the scores are confounded by differences in measurement and scaling properties that are group specific. The opposite is the case when there is support for measurement invariance.

A powerful method for examining measurement invariance is the multiple-group CFA mean and covariance structures (MACSA) approach. When the focus is on first-order factor models, this approach can test for configural invariance (same overall factor structure), item factor loadings invariance (same strength of the associations of items with the first-order factors), item intercepts (when the item scores are treated as continuous) or threshold (when the item scores are treated as ordered categorical) invariance (equivalence in item intercepts or threshold values), and error variances or uniqueness invariance (equivalence in the error variances of the items or variances of the items not attributed to the underlying constructs). When there is support for invariance for item factor loadings and intercepts or thresholds (as the case maybe), the groups can be also compared for their first-order latent factor mean scores [12]. Although the invariance

of the structural components (latent variances and covariances) can also be evaluated with MACSA, this evaluation is not relevant to measurement invariance.

For second-order factor models, the MACSA approach can be used to examine the invariance for the second-order configural model (same overall factor structure for the second-order factor model), and invariance for the second-order factor loadings (same strength of the associations of the primary factors with their secondary factors), first order factor intercepts (equivalence in intercepts values of the regression of the primary factors with the secondary factors), and first-order factor disturbances (same specific factors or unique variances for the primary factors that are not shared by the relevant common higher order factor). If there is support for the second-order factor loadings and the first order factor intercepts, then the groups can be compared for the second-order latent factor mean scores [13,14].

To date, at least three studies have examined invariance for CDI first-order factor models across children and adolescents [3,4,7]. For the child and adolescent models that they proposed, Weiss et al. found no support for configural equivalence and therefore measurement invariance. Garcia et al. examined invariance for the child and adolescent models proposed by Craighead et al. and their own models. For all these models they found support for equivalencies for the configural model, factor loadings model, and factor covariances, but not for item uniqueness. Scott et al. [7] found support for configural invariance and some items being non-invariant for factor loadings for a one-factor model. The differences in findings across these two studies may be related to the extraction procedures applied. The Weiss et al. study applied maximum likelihood (ML) extraction, whereas the Garcia et al. study used weighted least square (WLS) extraction. Scott

et al. [7] used robust WLS (WLSMV). Relative to ML, the WLS extraction is a more appropriate extraction procedure for analysis of categorical data. Evidence suggests that the application of ML extraction to categorical data, especially when there are four or fewer categories, like the CDI, will provide inaccurate and less accurate parameter estimates [15]. Although the WLS is suited for categorical data, it can led to substantial estimation difficulties with complex model, and accurate estimates need extremely large samples [10]. Clearly more studies in this area are needed, preferable with extractions procedures, such as the WLSMV that can minimize these difficulties [10]. Although this extraction method was applied Scott et al. [7], it was on a one-factor model, and not the original five-factor model proposed by Kovacs.

In addition to the existing contradictory findings, there are also limitations and omissions in the existing invariance data in this area. First, no study has tested invariance across children and adolescent for the original Kovacs [1] model. This can be seen as a significant omission as the Kovacs model is the model that is generally used in scoring. Indeed the scores and the scoring method provided in the CDI manual is based on the Kovacs model. Second, there has been no study of invariance for the second-order factor structure of the CDI across children and adolescents. Since the CDI scoring system is based on the total CDI score that is underpinned by the second-order factor model of the CDI, examination of the invariance for this level of this model is needed [16]. Third, to date there has been no study of invariance across clinic-referred children and clinic-referred adolescents. Given that the CDI is primarily used in clinical settings for screening the depressive disorders, such information will be clinically valuable and necessary. There are reasons to suspect that some non-invariance is possible as existing data show increaase in depression, as measured by the CDI, from childhood to adolescents [17,18].

Given the inconsistent findings, limitations and omissions in existing data, the first aim of the current study was to use CFA procedure appropriate for categorical data to examine support for the hierarchical CDI factor model proposed by Kovacs [1] for ratings provided by clinic-referred children and adolescents. We also tested the Kovacs's first order factor model by itself to allow comparisons with existing studies. Both sets of analyses were conducted for the sample as a whole, and for children and adolescents separately. Contingent on support for the hierarchical CFA factor, the second aim of the study was to use the MACSA approach that is appropriate for categorical data to examine support for measurement invariance for this model across clinic-referred children and adolescents. The third aim of the study was to compare the groups for the latent mean scores for the first-order latent factors (Negative Mood, Interpersonal Problems, Ineffectiveness, Anhedonia, and Negative Self-Esteem) and the second-order latent factor (General Depression).

Method

Participants

The data for all participants were collected archivally from the Academic Child Psychiatry Unit (ACPU) of the Royal Children's Hospital, Melbourne, Australia. The ACPU is an out-patient

psychiatric unit that provides services for children and adolescents with behavioural, emotional and learning problems. Only children and adolescents, between 7 and 17 years, who had completed the CDI were included in the study. In all, the data from 802 children and adolescents were included in this study. The participants in this study were the same ones as those used in a previous study that examined the measurement and factorial invariance of the CDI ratings for those with and without depressive disorders [19].

The participants were divided into separate child (N=459) and adolescent (N=343) groups. Like most previous studies in this area [1,4] children between 7 and 12 years were allocated to the child group, and those between 13 and 17 years were allocated to the adolescent group. The mean (SD) ages for the child and adolescent groups were 10.08 (1.39) years and 14.27 (1.16) years, respectively. The groups differed for age, t (800)=45.33, p<0.001, with large effect size, Cohen's d=3.23, based on Cohen's (1992) guidelines (small ≤ 0.20, medium ≥ 0.50, and large ≥ 0.80).

Demographic and background information for the child and adolescent groups are provided in **Table 1**. Mother and father employment status were recorded (and coded) as follows: Employed (1), home duties (2), pensioner (3), unemployed (4), student (5), other (6) and retired (7). Mother and father education (highest level) were recorded (and coded) as follows: Tertiary (7), high school or equivalent (6), technical certificate or equivalent (5), some years of secondary school (4), primary school (3), some years of primary school (2) and no schooling at all (1). The family income was coded as follows: $0-$30,000 (1), $30,000-$40,000 (2), $40,000-$50,000 (3) and $50,000 and over (4). **Table 1** shows the scores for these variables, treated as continuous. **Table 1** also shows the percentages of different groups of disorders for the child and adolescent groups, derived using the parent version of the Anxiety Disorders Interview Schedule for Children [20]. In the table, "any anxiety disorder" includes Separation Anxiety, Social Phobia, Specific Phobia, Panic, Agoraphobia, Generalized Anxiety, Obsessive Compulsive and/or Post-Traumatic Stress disorders. "Any depressive disorder" includes Dysthymic and/or Major Depressive Disorders.

As shown in **Table 1**, there were relatively more males than females in the child group, and more females than males in the adolescent group, with medium effect sizes in both cases. Although mothers of the child group had higher employment status than mothers of the adolescent group, the different was of small effect size. The groups did not differ for father's employment status, and mother's and father's educational levels. The frequency of depressive disorders was higher in the adolescent group, while the frequencies of other disorders were the same in child and adolescent groups. The effect size for the difference involving depressive disorders was medium.

Ethics

The study was approved by the RCH ethics committee as part of our group's comprehensive examination of children and adolescent referred for psychological problems. Each legal guardian and participant provided informed written consent

Table 1 Demographics in formation and disorders for the child and adolescent groups.

DSM-IV Diagnosis	Child	Adolescent	Test Statistics	Cohen's d
Age-Mean (SD)	10.08 (1.39)	14.27 (1.16)	t (df)=45.33***(800) χ² (df)=14.36*** (1)	3.23
Sex				
Boy N (%)	353 (76.9)	222 (64.7)	-	0.33
Girl N (%)	106 (23.1)	121 (35.2)		0.33
Mother's employment	1.91 (1.72)	1.72 (1.06)	t (df)=2.14* (782)	0.13
Mother's education	5.08 (1.27)	5.12 (1.30)	t (df)=0.41 (694)	0.03
Father's employment	1.83 (1.67)	1.78 (1.58)	t (df)=0.41 (694)	0.03
Father's education	5.11 (1.31)	4.94 (1.28)	t (df)=1.68 (681)	0.13
Family income	2.33 (1.32)	2.49 (1.35)	t (df)=1.63 (724)	0.12
DSM-IV Disorders				
Any ADHD	76.0	74.6	χ² (df)=0.11 (1)	0.04
CD/ODD	68.4	68.5	χ² (df)=0.19 (1)	0.00
Any anxiety disorder	73.0	70.6	χ² (df)=0.34 (1)	0.07
Any depressive disorder	41.4	56.3	χ² (df)=18.14** (1)	0.33

Note: ADHD: Attention Deficit/Hyperactivity Disorder; CD/ODD: Conduct Disorder/Oppositional.
*p<0.05, **p<0.01, ***p<0.001.

for any data provided by them to be used in future ethics approved research studies. This is a standard part of the ACPU assessment procedure. The RCH ethics committee adheres to the ethical guidelines set by the Australian National Medical Research Council. That in turn confirms to the World Medical Association Declaration of Helsinki - Ethical Principles for Medical Research Involving Human Subjects.

Measures

Children's Depression Inventory (CDI) [1]. As mentioned previously, the CDI is a self-rating scale for measuring depression in children and adolescents, aged 7-17 years. It can be administered individually or in groups. It has 27 items, and for each item, participants are asked to choose one of three statements that best describes them for the past 2 weeks. The options are graded in increasing level of clinical severity, from 0 to 2. For the current sample, the coefficient alpha values of the full scale were 0.88 for children and 0.90 for adolescents.

Procedure

The study had ethical approval from the Eastern Health and Royal Children's Hospital Ethics Review Boards and all participants' parents and children/adolescents gave informed consent for data collection. Children, adolescents and parents participated in separate interviews and testing sessions, with breaks, over a period of two days. Information was also obtained from teachers using various checklists and questionnaires. In all cases, parental consent forms were completed prior to the assessment. The data collected covered a comprehensive demographic, medical (primarily neurological and endocrinological), educational, psychological, familial and social assessment of the child and his or her family. All psychological data were collected by research assistants, who were advanced doctoral students in clinical psychology, and under the supervision of two registered clinical psychologists. The research assistants were provided with extensive

supervised training and practise by the two psychologists prior to them collecting data.

Statistical procedures

All the CFA models in the study were computed with Mplus (Version 6.1) software [21]. All the analyses used WLSMV. For evaluating model fit at the statistical level, the WLSMV estimation procedure produces the WLSMVχ². Like all other χ² values, this value is inflated by large sample sizes. Consequently, fit of the models was evaluate using the approximate (or practical) fit indexes of root mean squared error of approximation (RMSEA) and the comparative fit index (CFI). The guidelines suggested are that RMSEA values close to 0.06 or below be taken as good fit, 0.07 to 0.08 as moderate fit, >0.08 to 0.10 as marginal fit, and >0.10 as poor fit. For the CFI, values close to 0.95 or above are taken as indicating good fit, and values close to 0.90 and <0.95 are taken as acceptable fit [22,23]. Misfit was inferred if either one of these indices suggested a poor fit (that RMSEA values more than .08 and CFI values more than 0.90).

Multiple-group CFA measurement invariance for the second-order CDI model and the differences in the second- and first-order factor means scores was tested using the procedure demonstrated by Chen et al. and others [13,14], with some variation to account for ordered-categorical scores [24]. This essentially involves comparing progressively a series of nested invariance models. The procedure begins with the test of configural invariance of the second-order CDI model (M0). For this model, the pattern of fixed and free factor loadings of the first- and second-order factor loadings are specified for the groups, but the values of all parameters are not constrained equal across these groups. Following the computation of the configural invariance of the second-order CDI model (M0), the test for invariance for first-order factor loadings (M1) is conducted. In this model, M0 is revised so that the corresponding item loadings are constrained equal across the groups. The invariance for the first- and second-order factor loadings (M2) is tested next. In this model, M1 is

revised so that the corresponding first-order factor loadings are constrained equal across the groups. The invariance for the first- and second-order factor loadings and item thresholds (M3) is tested next. For this model, M2 is revised so that the corresponding item threshold values are constrained equal across the groups. Invariance for the first- and second-order factor loadings, item thresholds, and intercepts of the first-order factors (M4) is tested next. To test this model, M3 is revised to allow equality in corresponding first-order factor intercepts across the groups. The next model tested is invariance for disturbances of the first-order factor model (M5). For this model, M4 is revised so that the corresponding disturbances of the first-order factors are constrained equal across the groups. The final model tested is invariance for item uniqueness or error residual variances (M6). For this model, M5 is revised so that the corresponding item uniqueness values are constrained equal across the groups.

To test group differences for the first-order latent factor mean scores, invariance of the factor loadings and thresholds are imposed equal across the groups. In addition to these levels of invariance, invariance of the first order factor intercepts and item thresholds are imposed equal across these groups to test group difference for the second-order latent factor mean scores. As is required, in both models, the appropriate latent factor mean scores for one group is set to zero, while the latent mean scores for the other group is freely estimated. Thus the latent mean scores reflect relative differences between the groups.

Given that the difference in χ^2 is also inflated by large sample sizes, with trivial differences showing significance, researchers have pointed out that this test is too conservative or runs the risk of detecting invariance where no appreciable invariance exists [25]. The simulation study by Chen [26] suggested that a difference of -0.01 or more in the CFI value, together with an increase of 0.015 or more in the RMSEA value, can be taken as indications of lack of invariance. For this study, measurement invariance and also equivalence in latent mean scores were rejected if (a) there was inadequate fit for the invariance model, and (b) if the critical change values of both the RSMEA and the CFI were reached. Using these indices also allowed the application of the same standards for evaluating model fit and the differences in model fit.

Results

Missing data

With the WLSMV estimator, missing values are treated as pairwise missing, and the model is estimated only from observations with full records. However the percentage of missing data was trivial (0.5%) in our data set.

Fit for the CDI Hierarchical Model Proposed by Kovacs [1].

Table 2 shows the fit values of the hierarchical model for all participants together and for children and adolescents separately. As shown, for all analyses, the RMSEA values showed good fit, while the CFI indicated adequate fit. **Figure 1** shows the completely standardized estimates for the analysis involving all participants together. All factor loadings for the first-order factors were salient (<0.40, and ranging from 0.41 to 0.90) and significant (p<0.001). Also, all factor loadings of the first-order factors on the second-

Table 2 Fit Indices of the CDI model proposed by Kovacs [1].

Participants	χ^2	df	RMSEA (90% CI)	CFI
First-Order Factor Model				
All participants	1119.85	314	0.057 (0.053 - 0.060)	0.925
Children	664.86	314	0.049 (0.044 - 0.055)	0.923
Adolescents	736.61	314	0.063 (0.057 - 0.069)	0.939
Hierarchal Factor Model				
All participants	1196.074	319	0.059	0.918
Children	861.83	319	0.054	0.905
Adolescents	553.55	319	0.054	0.949

Note χ^2: Weighted least square with mean and variance adjusted chi-square (WLSMVχ^2); RMSEA: Root Mean Square Error of Approximation; CFI: Comparative Fit Index. All WLSMVχ^2 values were significant (p<0.001). CI values are not provided in MPlus output for Second-Order factor model.

order factor were salient (ranging from 0.72 to 0.94) and significant (p<0.001). Although not shown, the factor loadings for children and adolescents separately were similar to those found for both groups together. Taken together, these findings provide sufficient support for the CDI hierarchical model proposed by Kovacs [1].

Measurement Invariance for Kovacs's hierarchical factor model of the CDI

Table 3 shows the results of the multiple-groups invariance testing for the CDI hierarchical model proposed by Kovacs [1]. As shown, there was good fit for the configural model (M0) in terms of the RMSEA. The CFI value indicated acceptable fit. These values provide sufficient support for the configural invariance model. A review of **Table 3** shows that with the exception of the invariant item uniqueness model **(M6 in Table 3)**, the RMSEA and CFI values for all the other invariance models were at least adequate. For the item uniqueness invariant model, the RMSEA indicated adequate fit, whereas the CFI value indicated unacceptable fit. For all models that were compared, the differences in the RMSEA and CFI values were within the cut-off values used for accepting invariance (decrease of ≥ 0.01 in CFI values and an increase of ≥ 0.015 or the RMSEA values) (Figure 2). These findings suggest support for full measurement invariance (equivalency for all factor loadings, thresholds, and uniqueness).
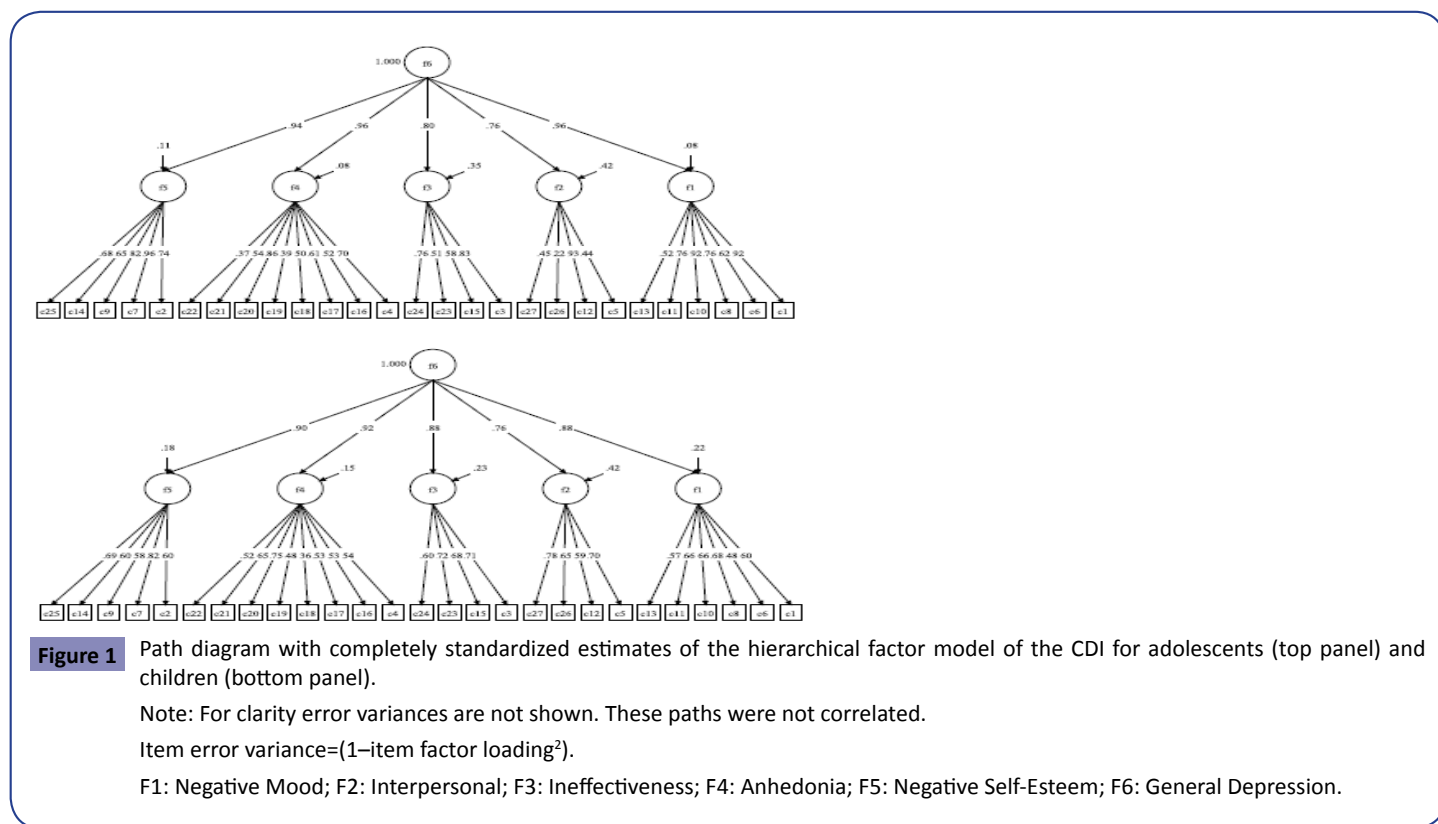
Group Differences for the First- and Second-Order Latent Mean Scores in Kovacs's Model

Given the invariance findings, the difference between the groups for the first- and second-order latent mean scores were examined. As shown in **Table 4**, for all five first-order latent factors, and for the second-order latent factors, adolescents scored higher (given that these values were positive and the values for children were set at zero). The effect sizes for the differences between the groups can be inferred by the standardized differences, which can be interpreted similarly to Cohen's guidelines [27,28]. The standardized differences are also presented in **Table 4**. As shown, the differences for the first-order factors for Negative Mood, Interpersonal Problems, Ineffectiveness, and Negative Self-Esteem were medium, while it was small for Anhedonia. The effect size for difference for the second-order latent factor or general depression was also medium.

Table 3 Results of tests for measurement invariance across children and adolescent ratings of the CDI.

	Model Fit				Model Difference		
	χ^2	df	RMSEA (90% CI)	CFI	ΔM	ΔRMSEA	ΔCFI
M0 (no constrains)	1457.00	640	0.056 (0.053 - 0.060)	0.930	-	-	-
M1 (invariant first-order factor loadings)	1564.63	662	0.058 (0.055 - 0.0620)	0.928	M1 – M2	0.002	-0.002
M2 (M1 plus invariant second-order factor loadings)	1578.96	666	0.058 (0.055 - 0.062)	0.922	M2 – M1	0.000	-0.000
M3 (M2 plus invariant item thresholds)	1660.52	686	0.060 (0.056 - 0.063)	0.917	M3 – M2	0.002	-0.005
M4 (M3 and invariant first-order factor intercepts)	1861.20	691	0.065 (0.062 - 0.069)	0.900	M4 – M3	0.005	-0.017
M5 (M4 and invariant first-order factor disturbances)	1861.62	696	0.065 (0.061 - 0.068)	0.901	M5 – M4	0.000	-0.001
M6 (M5 and invariant item uniquenesses)	1998.43	723	0.066 (0.063 - 0.070)	0.891	M6 – M5	0.001	-0.010

Note: χ^2: Weighted least square with mean and variance adjusted chi-square (WLSMVχ^2), RMSEA: Root mean Square Error of Approximation; CFI: Comparative Fit Index. All WLSMVχ^2 values were significant (p<0.001).

Figure 1 Path diagram with completely standardized estimates of the hierarchical factor model of the CDI for adolescents (top panel) and children (bottom panel).

Note: For clarity error variances are not shown. These paths were not correlated.

Item error variance=(1–item factor loading^2).

F1: Negative Mood; F2: Interpersonal; F3: Ineffectiveness; F4: Anhedonia; F5: Negative Self-Esteem; F6: General Depression.

Discussion and Conclusion

The results of the study indicated support for the first-order CDI factor models proposed by Kovacs [1]. These results were found for all participants together, and for children and adolescents separately. Unlike this study, the CFA study by Garcia et al. [3] failed to find support for this model for both children and adolescents. These discrepant findings may be related to differences in the type of samples examined and extraction procedure applied in the CFA. This study examined a clinic-referred sample and applied WLSMV, whereas Garcia et al. examined a community sample and applied WLS. The findings here also found support for Kovacs's hierarchical model. This is the first study to use CFA to directly test and find support for this model. Taken together the findings in this study indicate that the original hierarchical model proposed by Kovacs is an acceptable theoretical model for the CDI, at least for ratings provided by clinic-referred children and adolescents.

The findings here also indicated support for measurement invariance for Kovacs's [1] hierarchical model across self-ratings from clinic-referred children and adolescents. More specifically, all first-order factor loadings (M1), second-order factor loadings (M2), item thresholds (M3), first-order factor intercepts (M4), and first-order factor disturbances (M5) were equivalent across these groups. There was mixed support for item uniqueness (M6). Since the test for equivalence in item uniqueness is generally considered stringent and of little substantive value in equivalence testing [13,29], the mixed support for this level of invariance is not problematic as such. Thus the findings in this study can be taken to mean that clinic-referred children and adolescents ratings of the CDI items, modelled in terms of Kovacs's [1] hierarchical model, have the same measurement and scaling qualities. It is worth noting that this is the first study to test and find support for measurement invariance for this model. Garcia et al. [3] also

Table 4 Results of the tests for differences in latent mean scores.

	Model Fit				Mean Difference for Latent Factors	
	χ^2	df	RMSEA (90% CI)	CFI	Estimate (Standardized)	z-value
First-Order Factors	1592.46	672	0.058 (0.055 - 0.062)	0.922	-	-
Negative Mood	-	-	-	-	0.437 (0.488)	5.02***
Interpersonal Problems	-	-	-	-	0.247 (0.400)	4.45*
Ineffectiveness	-	-	-	-	0.512 (0.612)	5.57***
Anhedonia	-	-	-	-	0.161 (0.245)	2.51*
Negative Self-Esteem	-	-	-	-	0.359 (0.425)	4.35***
Second-Order Factor	1696.14	694	0.060 (0.056 - 0.064)	0.915	-	-
Depression	-	-	-	-	0.372 (0.444)	4.80***

Note: χ^2: Weighted least square with mean and variance adjusted chi-square (WLSMVχ^2), RMSEA: Root Mean Square Error of Approximation; CFI: Comparative Fit Index. All WLSMVχ^2 values were significant ($p<0.001$).

*$p<0.05$, **$p<0.01$, ***$p<0.001$.

found support for measurement invariance across these groups from the general community for model they proposed and for Craighead et al. [6] models.

The support for measurement invariance for the hierarchical factor model proposed by Kovacs [1] has important implications for the clinical use of the CDI. The support suggests that the ratings and observed scores provided by clinic-referred children and adolescents can be justifiably compared as they are not confounded by different measurement or scaling properties. This means that the same response categories will be endorsed by these groups when they have the same levels of underlying depression. Also, as the total score is underpinned by the hierarchical factor, the support for invariance for this model means that the total scores from these groups are also directly comparable. This is valuable information as the total score is computed and used in the same way for screening depressive disorders in these groups.

The invariance findings found in this study have implications for understanding if the developmental changes in 'depression' reflect "heterotypic continuity" or "homotypic continuity". Heterotypic continuity suggests that there are developmental differences in how depressive symptoms are expressed, but the symptoms do not differ when considered as higher level constructs. Homotypic continuity suggests phenotypic or symptomatic consistency across development. While some researchers have argued in favor of the heterotypic continuity argument [30,31] have noted that the general consensus is that the essential symptoms for 'depression' have homotypic continuity. The support here for full measurement invariance means that developmental levels have relatively little influence on the phenomenology of depressive symptoms (at least during and the childhood and adolescent periods), and is therefore consistent with the homotypic continuity argument.

The results of this study also showed that adolescents had higher scores than children for both the first- and second-order latent factors. The effect sizes for the first-order factors for Negative Mood, Interpersonal Problems, Ineffectiveness, and Negative Self-Esteem were medium, while it was low for Anhedonia. The effect

size for difference for the second-order latent factor or general depression was also medium. These findings suggest that despite homotypic continuity of depression symptoms from childhood to adolescence, clinic-referred adolescents can be expected to express moderately more severe levels of these symptoms than clinic-referred children. Our findings are consistent with existing CDI data [17,18], and also with the view that depression increases noticeably among adolescents (in particular among females) following the onset of puberty [32-34]-a finding also reported specifically for the CDI [18].

In concluding, it needs to be noted that the findings and interpretations made in the study need to be viewed with some limitations in mind. First, the findings reported here are based on a single study. As a consequence, there is a need for cross-validation of the findings before the findings can be generalized. Second, parental concerns may vary across children and adolescents leading to different reasons for referral for these developmental groups [35], which in turn could influence observed developmental differences. Third, all the participants in this study were from the same clinic. Thus it is possible that this may constitute an additional bias for the sample examined, limiting the findings and conclusions made in this study. Fourth, as this study was on clinic-referred children and adolescents, the applicability of the findings for children and adolescents in the general community cannot be assumed. Fifthly, in the invariance tests, nested models were compared using the differences in two approximate fit indexes (RMSEA and CFI). Thus the invariance findings are best view from a practical viewpoint, and cannot to be viewed from a statistical viewpoint. It will be useful for future studies to examine samples from several clinics and from the general community in the same study, In the meantime it is worth noting that the findings in the current study indicate support for the original CDI model proposed by Kovacs, and also that the ratings provided by clinic-referred children and adolescents interpreted in terms of this model can be compared as they are not confounded by different measurement and scaling properties. Thus the CDI as proposed originally by Kovacs had sound utility for clinical use with clinic-referred children and adolescents.

References

1   Kovacs M (1992) Children's Depression Inventory. New York: Multi-Health Systems.

2   Sun S, Wang S (2015) The Children's Depression Inventory in worldwide child development research: A reliability generalization study. J Child Fam Stud 24: 2352-2363.

3   Garcia LF, Aluja A, del Barrio V (2008) Testing the hierarchical structure of the Children's Depression Inventory: A multigroup analysis. Assessment 15: 153-164.

4   Weiss B, Weisz JR, Politano M, Carey M, Nelson WM, et al. (1991) Developmental differences in the factor structure of the Children's Depression Inventory. Psychol Assessment 3: 38-45.

5   Logan DE, Claar RL, Guite JW, Kashikar-Zuck S, Lynch-Jordan A, et al. (2013) Factor Structure of the Children's Depression Inventory in a multisite sample of children and adolescents with chronic pain. J Pain 14: 689-698.

6   Craighead WE, Smucker MR, Craighead LW, Ilardi SS (1998) Factor analysis of the Children's Depression Inventory in a community sample. Psychol Assessment 10: 156-165.

7   Scott, WD, Clapp J, Mileviciute I, Mousseau A (2016) Children's Depression Inventory: A unidimensional factor structure for American Indian and Alaskan Native youth. Psychol Assessment 28: 81-91.

8   Taylor JJ, Grant KE, Amrhein K, Carter JS, Farahmand F, et al. (2014) The manifestation of depression in the context of urban poverty: A factor analysis of the Children's Depression Inventory in low-income urban youth. Psychol Assessment 26: 1317-1332.

9   Steele RG, Little TD, Ilardi SS, Forehand R, Brody GH, et al. (2008) A confirmatory comparison of the factor structure of the Children's Depression Inventory between European American and African American youth. J Child Fam Stud 15: 779-794.

10 Brown TA (2006) Confirmatory factor analysis for applied research. New York: Guilford Press.

11 Reise SP, Widaman KF, Paugh PH (1993) Confirmatory factor analysis and item response theory: two approaches for exploring measurement invariance. Psychol Bull 114: 552-566.

12 Widaman KF, Reise SP (1997) Exploring the measurement invariance of psychological instruments: Applications in the substance use domain. In: Bryant KJ, Windle M, West SG (eds.) The science of prevention: Methodological advances from alcohol and substance abuse research. Washington, DC: American Psychological Association, pp: 281-324.

13 Byrne BM, Stewart SM (2006) The MACS approach to testing for multiple group invariance of a second-order structure: a walk through the process. Struct Equ Modeling 13: 287-321.

14 Chen FF, Sousa KH, West SG (2005) Testing measurement invariance of second-order factor models. Struct Equ Modeling 12: 471-492.

15 Finney SJ, DiStefano C (2006) Non-normal amd categorical data in structural equation modelling. In: Hancock GR, Mueller RO (eds.) Structural Equation Modeling: A Second Course. Greenwich, Connecticut, Information Age Publishing.

16 Carle AC, Millsap RE, Cole DA (2008) Measurement bias across gender on the Children's Depression Inventory: Evidence for invariance from two latent variable models. Educ Psychol Meas 68: 281-303.

17 Twenge JM, Nolen HS (2002) Age, gender, race, socioeconomic status, and birth cohort differences on the Children's Depression Inventory: A meta-analysis. J Abnorm Psychol 111: 578-588.

18 Roelofs K, Braet C, Rood L, Timbremont B, van Vlierberghe L, et al. (2010) Norms and screening utility of the Dutch version of the Children's Depression Inventory in clinical and nonclinical youths. Psychol Assessment 22: 866-877.

19 Gomez R, Vance A, Gomez A (2012) Children's Depression Inventory: Invariance across children and adolescents with and without depressive disorders. Psychol Assessment 1: 1-10.

20 Silverman WK, Albano AM (1996) Manual for the ADIS-IV C/P. New York: Psychological Corporation.

21 Muthen LK, Muthen BO (2006) Mplus user's guide (4th edn) Los Angeles, CA: Muthen & Muthen.

22 Hu LT, Bentler PM (1998) Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. Psychol Method 3: 424-453.

23 Steiger JH (1998) A note on multiple sample extensions of the RMSEA fit index. Structural Equation Modeling 5: 411-419.

24 Millsap RE, Yun-Tein J (2004) Assessing factorial invariance in ordered-categorical measures. Multivariate Behav Res 39: 479-511.

25 Cheung GW, Rensvold RB (2002) Evaluating goodness-of-fit indexes for testing measurement invariance. Struct Equ Modeling 9: 233-255.

26 Chen FF (2007) Sensitivity of goodness of fit indices to lack of measurement invariance. Struct Equ Modeling 14: 464-503.

27 Cohen J (1992) A power primer. Psychol Bull 112: 155-159.

28 Thompson M, Green SB (2006) Evaluating between group differences in latent variable means. In: Hancock GR and Mueller RO (eds.) Structural equation modelimg: A second course, Greenwich, CT: Information age publishing, pp: 119-169.

29 Cheung GW (2008) Testing equivalency in the structure, means, and variances of higher-order constructs with structural equation modelling. Organ Res Meth 11: 593-613.

30 Carlson GA, Kashani JH (1988) Phenomenology of major depression from childhood through adult-hood: Analysis of three studies. Am J Psychiatry 145: 1222-1225.

31 Weiss B, Garber J (2003) Developmental differences in the phenomenology of depression. Dev Psychopathol 15: 403-430.

32 Angold A, Costello EJ, Erkanli A (1999) Comorbidity. J Child Psychol Psychiatry 40: 57-87.

33 Birmaher B, Ryan ND, Williamson DE, Brent DA, Kaufman J, et al. (1996) Childhood and adolescent depression: A review of the past 10 years. Part I. J Am Acad Child Adolesc Psychiatry 35: 1427-1439.

34 Kessler RC, McGonagle KA, Swartz M, Blazer DG, Nelson CB (1993) Sex and depression in the National Comorbidity Survey. I: Lifetime prevalence, chronicity and recurrence. J Affect Disord 29: 85-96.

35 Gottlieb J, Gottlieb BW, Trongone S (1991) Parent and teacher referrals for a psychoeducational evaluation. J Spec Educ 25: 155-167.