Psychologica Belgica
2013, 53/2, 51-74

ADEQUATE SCREENING OF YOUNGSTERS FOR DEPRESSIVE CHARACTERISTICS

Introduction. In order to set up an effective early-detection of depressive symptoms in youngsters, the current study aims to investigate whether two measure moments of the Children's Depression Inventory (CDI) improve screening and whether a multi-informant procedure is superior compared to a single-informant procedure thereby controlling for comorbid symptoms. **Method.** Youngsters (10-15 years) filled in the CDI and an Anxiety Scale at Time 1 and the CDI and Youth Self Report one week later. Next, a structured clinical interview was administered. The Child Behaviour CheckList was filled in by the parents. **Results.** Two measure moments of the CDI are not more accurate in capturing disordered mood changes. Furthermore, parent reports were no significant contributor to the variance over and above the CDI. **Discussion.** A second moment does not increase screening accuracy. Further research on setting up an effective multistage screening procedure for depressive symptoms for youngsters is however necessary.

Method

Depression in children and adolescents is a severe disorder that should be detected and treated early in its development (Zwaanswijk, Verhaak, Bensing, van der Ende, & Verhulst, 2003). Early identification may create opportunities to reduce the impact of depression on the child and, more specifically, early intervention may improve social and academic functioning and lower risks of suicide, substance abuse, and persistence of depressive disorders into adulthood (Birmaher, Brent, & AACAP Workgroup on Quality Issues, 2007). Consequently, the American Academy of Child and Adolescent Psychiatry has proposed a practice parameter based on research and current clinical practice (Birmaher et al., 2007). They recommend that practitioners routinely include screening questions about depressive symptoms in the psychiatric assessment of children and adolescents. The present investigation aims to address this call by evaluating a screening procedure for depression in children.

* Lotte Theuwis, (MD), Caroline Braet (PhD), & Eva Vandevivere (MD), Department of Developmental, Personality, and Social Psychology, Ghent University, Ghent, Belgium; Jeffrey Roelofs (PhD), Department of Clinical Psychological Science, Maastricht University, Maastricht, The Netherlands; Kevin Stark (PhD), Department of Educational Psychology, University of Texas, United States.
Correspondence concerning this article should be addressed to Caroline Braet, Department of Developmental, Personality, and Social Psychology, Henri Dunantlaan 2, 9000 Ghent, Belgium. E-mail: caroline.braet@Ugent.be

In 2002 the US Preventive Services Task Force (USPSTF) concluded that there was unsatisfactory evidence to recommend for or against routine screening of youngsters for depression. One of the concerns they put forward was the lack of sufficient psychometric data on the available instruments. In 2009 the Oregon Evidence-Based Practice Center conducted a review of the existing evidence on screening for Major Depressive Disorder (MDD) among children and adolescents from primary care populations (Williams, O'Connor, Eder, & Whitlock, 2009). Two conclusions were of importance here. First, most studies reported poor construct validity since many of the rating scales measured general distress instead of depressive symptoms. Second, their ten-year review concluded also that specifically data on the accuracy of depression screening instruments in *younger* children remains limited. Most studies restricted study samples to adolescents aged 12 years or older and they administered self-report measures originally designed for adults. Although self-report questionnaires have the advantage of providing insight into the subjective perception of a person (Garber & Kaminski, 2000), some researchers question their reliability in younger age groups (Birmaher et al., 2007). Sensitivity rates in these samples appear to be remarkably lower (Garrison, Addy, Jackson, McKeown, & Waller, 1991; Garrison, Jackson, Marsteller, McKeown, & Addy, 1990; Goodman, Ford, Simmons, Gatward, & Meltzer, 2003; in Williams et al., 2009). This demonstrates the need for research on screening for depressive symptoms in younger age groups using appropriate instruments.

Previous research has demonstrated that depressive disorders among children represent a heterogeneous phenomenon (Harrington, Rutter & Fombonne, 1996). The international literature argues for the phenomenological equality of depressive symptoms across all ages instead of emphasising the differences (Birmaher et al., 2007). On the other hand child psychologists are still questioning how to assess depression during childhood because the general definition fails to recognise the developmental aspects of the disorder. Overall, the "core symptoms" of MDD in children and adolescents (such as sad or irritable mood and decreased interest, pleasure, or motivation) are similar to those in adults. However, the physical, emotional, cognitive and social developmental stages of children will have different effects on the manifestation of the peripheral symptoms (Klein, Dougherty, & Olino, 2005; Yorbik, Birmaher, Axelson, Williamson, & Ryan, 2004). This differential manifestation could influence the sensitivity rates of screening instruments originally created for adults. Therefore it is important to work with age-adapted screening instruments.

Some instruments adapted for children and adolescents have already proved to be useful. For example, the Children's Depression Inventory (CDI), (Kovacs, 1992) has demonstrated acceptable psychometric properties. The

CDI has high levels of internal consistency, test-retest reliability, and predictive, convergent, and construct validity (Craighead, Smucker, Craighead, & Llardi, 1998; Timbremont & Braet, 2002). The sensitivity and specificity of the CDI compared with diagnoses derived from a structured clinical interview has been assessed in a clinical (Timbremont, Braet, & Dreesen, 2004) and by Stark in a school sample (1990) and guide the use of clinical cut-offs of the CDI. A cut-off score of 16 showed an optimal relation between sensitivity (94.4%) and specificity (83.8%) (Timbremont, et.al., 2004, Stark, 1990). The CDI appears to be an eligible candidate as a screening instrument in younger children and adolescents. However, some issues still need to be researched more profoundly.

First, irritable mood is an important and prominent feature of depression in children and adolescents. However, irritability appears more frequently in youngsters compared to adults and typically occurs during puberty. This makes it difficult to distinguish from moodiness as a characteristic from depression. In essence, bursts of irritability are seen as normal emotional responses in puberty, whereas frequent and persistent irritability may be a symptom of a wide variety of disorders. These normal variations of moodiness may result in false positives when screening for depressive symptoms in adolescents since one-wave assessments are commonly used in research and clinical practice. Roberts, Lewinsohn and Seeley (1991) brought to light that a repeated administration of one and the same rating scale in an adolescent sample improved the discriminant validity. Nelson and Politano (1990) compared repeated administration of the CDI in a sample of 96 psychiatrically hospitalised children 6 to 15 years of age. They found that respondents endorse greater symptomatology on the first administration and that initial scores significantly decreased on the second and third administrations. Therefore, extent literature calls into question the method of relying on one single administration of a self-report instrument when screening for depression in children. Reynolds already stated in 1986 that depression, even in adults, is not a stable trait and therefore multistage approach, with multiple measure moments, is supposed to be superior to a single-moment assessment. Especially in children it is important to investigate whether a single stressor or specific event is the cause of a transient mood fluctuation or if scores are the result of an ongoing disorder. A second screening moment can refine the screening process and eliminate children who either dissemble or exaggerate their reports of depressive symptoms, experience a transient depressed mood, or for other reasons score high on one measure moment but below the cut-off on a second one (Reynolds, 1986).

Second, due to the subjective nature of internalising disorders self-reports are often assumed to be the only important perspective. Still, also a multimethod procedure has been recommended where the administration of a self-

report measure is followed by administration of a clinical interview at the end of this stage (Reynolds, 1986). Because of incomplete language development and immature cognitive processes in youngsters the reliability of self-reports is seen as a concern (Ollendick, Grills, & King, 2001; Zeman, Klimes-Dougan, Cassano, & Adrian, 2007). Therefore other researchers have suggested that this issue could be addressed through the use of multiple information sources (Kovacs, 1986). However, evidence-based guidelines have not yet been established for this practice.

Generally, when compared to clinicians' judgements an underestimation of real depressive disorders is noticed in the family environment because internalising disorders are more difficult to detect and, in addition, moodiness may initially be seen as age-appropriate (Lagges & Dunn, 2003). Still there is convincing evidence for the use of both parent and child reports (Jensen et al., 1999). Although it is recognised that, due to their economic quality self reports, parental reports, and teacher rating scales can all be used as screening instruments (Birmaher et al., 2007), it is still puzzling how the so-called multi-method multi-informant procedure must be integrated to get an early detection of depressed children. To extend the existing evidence we will specifically focus on children ages 10 to 15 years old. Our research questions will be evaluated in a school sample since investigating this target group is the most adequate way to evaluate the effectiveness of early detection of severe depressive symptoms in youngsters. Only when testing the specific use of a screenings procedure in a general population group we can find out how routine screening must be designed.

Some investigators state that heightened scores on a depression rating scale could be the sign of general distress instead of a specific indication of mood problems (e.g. Fendrich, Weissman, Warner, 1990). For that reason, we will control whether youngsters who score high on the CDI are depressed and not suffering from other psychopathology like anxiety or externalising behaviour. A self-report questionnaire measuring anxiety symptoms and a broadband screening instrument for externalising psychopathology, to control for comorbid psychopathology, will be included in the procedure.

The *first aim* of this investigation is to evaluate whether two measure moments of a depression rating scale, CDI (Kovacs, 1992) contribute to a better explanation of variance of depressive symptoms and diagnosis of depressive disorders relative to a single measurement. Or in other words, to determine whether two measure moments more reliably capture disordered moods compared to one. We expect that two assessments will provide a more reliable and stable picture of depressive symptoms, instead of a single indication of the sad mood at a particular moment.

Second, we will compare a multi-informant procedure (child & parent reports on the ASEBA internalising scales; Achenbach & Rescorla, 2001) to

a single-informant procedure using the KID-SCID (Structured Clinical Interview for DSM-IV Child edition; Hien et al., 1994) as a gold standard. We will test whether combining both information sources improves detection of depressive symptoms, without masking distinctive patterns and specific diagnostic information related to the point of view by which one is reporting. So, conform the literature which shows substantial disagreement between reports of multiple informants (children & parents), we predict that both will account for a significant and unique portion of the variance in predicting clinically relevant diagnoses (Ferdinand et al., 2003; Verhulst, Dekker, & van der Ende, 1997) and hereby will improve the discriminant validity of the assessment process.

Third, in all these analyses, we will control for comorbid psychopathology and check whether heightened scores on the CDI are explained by depressed mood and not by other disorders. We expect that the results of this study could serve as a starting point for designing a screening strategy for identification of depression among children and adolescents in a school environment.

Method

Procedure

The study protocol was approved by the Ethical Commission for Scientific Research of our institute. A total number of 400 youngsters of the primary and secondary education were invited to participate. Informed consent was obtained from the participating schools, the youngsters and their parents. Eventually, 347 youngsters, and their parents agreed to participate, which corresponds with a response rate of 86,8%. The youngsters did not receive any credit in return for participating.

Following clinical recommendations for a multistage strategy, we used three moments of assessment (Kendall, Cantwell, & Kazdin, 1989). In the first stage youngsters were asked to complete the Children's Depression Inventory (CDI-T1), (Kovacs, 1992) and the Revised Child Anxiety and Depression Scale (RCADS), (Chorpita, Yim, Moffitt, Umemoto, & Francis, 2000; Muris, Meesters, & Schouten, 2002). Exactly one week later (Time 2) the youngsters were asked to complete the CDI a second time (CDI-T2) and the Youth Self Report (YSR), (Achenbach & Rescorla, 2001; Dutch version by Verhulst & van der Ende, 2004). A research assistant was available on both assessment moments to answer questions and to ensure the confidentiality and independent responding of participants. The parent who spent most of the time with the child was asked to fill in the Child Behaviour CheckList (CBCL), (Achenbach & Rescorla, 2001; Dutch version by Verhulst & van der Ende, 2004) at home (between time 1 & 2). At Time 3, one month after Time

1, the Structured Clinical Interview for DSM-IV Child Edition (KID-SCID), (Hien et al., 1994; Dutch version by Dreessen, Stroux, & Weckx, 1998) was administered by a research assistant who was trained in standardised administration of the KID-SCID.

Since it was too time-consuming to interview all of the participants, a decision rule was followed for administering the KID-SCID. Participants who scored above 10 on either the time one or time two on CDI were interviewed with the KID-SCID. To control for false negatives an equal number of participants who scored below 10 were randomly chosen to complete the KID-SCID interview. We chose a lower cut-off point than originally recommended by Kovacs (1992) to assure that children with fewer symptoms were also further interviewed.

All youngsters between 10 and 15 years old following regular education were eligible for this participation. School sampling was based on grade and type of curriculum, so the sample contains youngsters from the last year of primary and first grade of secondary education. In total we contacted 9 schools situated all over Flanders, of which 7 agreed to participate. In each school we randomly picked out one class of each educational type and grade.

Parents of participants who were diagnosed with a depressive disorder or appeared to be at the start of developing a depressive disorder were contacted and offered the opportunity to have their child enter a 16-week treatment protocol for depression at no cost to the family.

Participants

Participants were 347 youngsters with a mean age of 12.42 years (SD = 0.85 years; range 10-15 years). The sample was composed of 181 boys (51.9%) and 168 girls (48.1%). The sample contained both students from primary schools (children; 12.9%) as well as of the first grade of secondary school (pre-adolescents) of which 42.1% followed general, 22.9% technical and 13.8% vocational training education (8.3% missing). The socioeconomic status (SES) was calculated using parents' education and current profession (Hollingshead, 1975). The sample was primarily middle class (47.5%), 15.90% was upper-middle to upper class, and 22.10% were lower-middle to lower class (14.5% missing values) which is representative according to the current SES distribution in this region.

Based on their total score on the CDI (> 10) 103 youngsters were eligible to participate in the KID-SCID interview and 64 youngsters (CDI < 10) were randomly chosen to participate, creating a final study sample of N = 167.

Measures

Structured Clinical Interview for DSM-IV Child edition (KID-SCID). The KID-SCID interview (Hien et al., 1994; Dutch version by Dreessen, et al., 1998) is based on the SCID for adults (Spitzer, Williams, & Gibbon, 1986), a widely used diagnostic interview that has acceptable reliability and validity (Spitzer, Williams, Gibbon, & First, 1992; Williams, Gibbon, First, & Spitzer, 1992). Similar to the adult version, the KID-SCID is a semi-structured instrument designed to generate childhood *DSM-IV* diagnoses for clinical research studies. In this study, the following modules were assessed: disruptive behaviour disorders, mood disorders, and anxiety disorders. The interview score is an investigator-based score whereby the answers of the child or adolescent were used as the information source. Pilot data of Matzner (1994) indicated excellent inter-rater reliability for the disruptive behaviour module (.84 for oppositional defiant disorder and conduct disorder and 1.0 for attention deficit hyperactivity disorder). In a Belgian sample Cohen's Kappa ranged from .79 to 1.0, suggesting excellent agreement (Roelofs et al., 2010).

All interviews were audiotaped and coded twice by different raters to evaluate the inter-rater reliability of the KID-SCID. Sixty-two interviews could not be included in the calculation of Cohen's Kappa because they had zero variance on all items. The inter-rater reliability rates are presented in Table 1 and reveal substantial to outstanding agreement (Landis & Koch, 1977).

Table 1
Inter-rater reliability of the KID-SCID interviews (N=105)

Diagnosis	Cohen's Kappa	Percentage agreement
Disruptive Disorders		
ADHD	.79	98%
ODD	1	100%
CD	/	
Mood Disorders		
Current Depressive Episode	.88	99%
Past Depressive Episode	.32	96%
Current (Hypo)Manic Episode	/	
Past (Hypo)Manic Episode	/	
Current Dysthymic Disorder	1	100%
Depressive Disorder	.73	96%
Depressive Disorder NOS	.66	99%
Anxiety Disorders		
Separation Anxiety Disorder	/	
Social Phobia	/	

Table 1
Inter-rater reliability of the KID-SCID interviews (N=105) (continued)

Diagnosis	Cohen's Kappa	Percentage agreement
Specific Phobia	.66	99%
OCD	/	
PTSD	/	
GAD	/	
Panic Disorder	/	
Agoraphobia	/	
Mean score over all disorders:	.76	98%

Note: ADHD = Attention Deficit Hyperactivity Disorder, ODD = Oppositional Defiant Disorder, CD = Conduct Disorder, OCD = Obsessive-Compulsive Disorder, PTSD = Post Traumatic Stress Disorder, GAD = General Anxiety Disorder
/ = No measures of association could be computed because at least one variable was a constant
Percentage agreement = calculated by summing the total number of agreements on diagnoses and no diagnoses, divided by total number of judgements both encoders have passed

In the present study the following alpha values were found for internal consistency: mood disorders (.96 for Current Depressive Episode, .94 for Past Depressive Episode, .95 for Major Depressive Disorder and .97 for Dysthymic Disorder and no alpha-value could be calculated for Bipolar Disorder since the scale had zero variance items), anxiety disorders (.76 for Separation Anxiety Disorder, .85 for Specific Phobia, .95 for PTSD, .95 for GAD, .87 for Panic Disorder, .72 for Agoraphobia and no alpha-value could be calculated for Social Phobia and OCD since the scale had zero variance items), disruptive behaviour disorders (.93 for ADHD, .90 for ODD and .68 for CD).

Children's Depression Inventory (CDI). The youngsters completed a Dutch version of the CDI (Kovacs, 1992; Dutch version by Timbremont & Braet, 2002). A back-translation procedure was used and acknowledged by Maria Kovacs. The Dutch version of the CDI directly corresponds to the original version, which is the child version of the Beck Depression Inventory (BDI), (Beck, Ward, Mendelson, Mock, & Erbaugh, 1961). The CDI is a self-report questionnaire used for youngsters ages 7 to 17 years and includes 27 items assessing cognitive, affective, and behavioural symptoms of depression. Each item consists of three statements graded in order of increasing severity from 0 to 2; youngsters select the one that characterised them best during the past 2 weeks. The item scores are combined into a total depression score. As already mentioned, the questionnaire has relatively high levels of internal consistency, test-retest reliability, and predictive, convergent, and construct validity, especially in nonclinical populations (Craighead et al., 1998; Timbremont & Braet, 2002). In the current sample Chronbach's alpha was 0.82 at Time 1 and .88 at Time 2.

The questionnaire can be interpreted by means of cut-off scores based on the raw total score. A cut-off score of 13 minimises false negatives and is recommended in a clinical sample; a higher cut-off score of 19 minimises false positives and can be used in a nonclinical sample (Kovacs, 1992). The screening boundary of 19 in non-clinical samples was confirmed by Stark and his colleagues (Stark, Humphrey, Laurent, Livingston, & Christopher, 1993). Until now two Dutch studies [15, 41] investigated the quality of the Dutch version of the CDI. In both studies a cut-off score of 16 maximised the specificity and sensitivity of the CDI and resulted in a total predictive value (percentage of cases correctly classified in depressed versus non-depressed) between 86.3% (Timbremont et al., 2004) and 93.23% (Roelofs, et al., 2010).

Revised Child Anxiety and Depression Scale (RCADS). The RCADS is an adaptation of the self-report questionnaire Spence Children's Anxiety Scale (SCAS; [46]) and designed to assess symptoms of DSM-defined anxiety disorders and major depression. The scale consists of 47 items that on the basis of exploratory factor analysis (Chorpita et al., 2000) are allocated to six subscales: social phobia (9 items), panic disorder (9 items), major depressive disorder (10 items), separation anxiety disorder (7 items), generalised anxiety disorder (6 items), and obsessive-compulsive disorder (6 items). Items have to be scored on a 4-point scale with 0 = never, 1 = sometimes, 2 = often, and 3 = always. RCADS subscale scores can be obtained by summing across relevant items. In this study the subscale "major depressive disorder" was not included in the analyses since there was too much overlap with the CDI. Chronbach's alpha reaches high levels of internal consistency in this sample (.86 for panic disorder; .83 for separation anxiety disorder; .85 for generalised anxiety disorder and .81 for obsessive-compulsive disorder) except for the scale "social phobia" disorder (.01). The Chronbach's alpha for "social phobia" decreases up to .86 when we delete the item "I'm worried about my looking foolish". In the analyses we will use the adjusted version of the subscale "social phobia", with the disturbing item deleted.

Child Behaviour CheckList (CBCL) & Youth Self Report (YSR). The CBCL and the YSR (Achenbach & Rescorla, 2001; Verhulst & Van der Ende, 2004) are parallel questionnaires assessing several emotional and behavioural problem areas as reported respectively by the parent and the child. For both informants, a global internalising, externalising and total problem raw score can be obtained. Dutch versions of both the CBCL and the YSR are reliable and valid instruments for the assessment of psychological symptoms in youth (Verhulst & Van der Ende, 2004). In the current study following alpha-values were computed for YSR internalising problem score ($\alpha = .85$), YSR externalising problem score ($\alpha = .89$), YSR total problem score ($\alpha = .93$), CBCL internalising problem score ($\alpha = .89$), CBCL externalising problem score ($\alpha = .87$) and CBCL total problem score ($\alpha = .94$).

Data Analyses

All data were analysed using SPSS 15.0 for Windows. Comparison of means and covariances of all questionnaire variables using Little (1988) MCAR test revealed that data were missing completely at random, χ^2 (207) = 214,78, p = .34. Therefore, we decided to estimate missing values using maximum likelihood estimation and the expectation maximisation algorithm available in SPSS (Schafer, 1997). Furthermore, we assessed the collinearity between the two measure moments of the CDI by examining the tolerance and Variation Inflation Factor (VIF). Both the levels of the tolerance (.40 for CDI-T1 and CDI-T2) as well as the VIF (2.53 for CDI-T1 and CDI-T2) were completely acceptable since a tolerance of less than 0.20 or 0.10 and/or a VIF of 5 or 10 and above indicates a multi-collinearity problem (Menard, 1995 in O'Brien 2007).

Based on scores on the KID-SCID we divided the final sample (N = 167) in two groups: (1) Current Depressive Symptoms (CDS) including all youngsters who were diagnosed with a Current Major Depressive Disorder, a Current Dysthymic Disorder or Current Depressive Disorder NOS*;* n = 14) and (2) all youngsters not receiving a diagnosis, here defined as No Current Depressive Symptoms (No-CDS; n = 153). This division is based on the DSM-IV criteria for a Current Episode of Depressive Disorder in children and adolescents. No children in the CDI < 10 group were detected as cases of CDS with the KID-SCID while in the CDI > 10 group, 14 cases of CDS were found. Frequencies of other diagnoses are reported in Table 2.

Table 2

Frequencies of diagnoses based on the KID-SCID interview (N = 167)

Diagnosis	N	Percentage
Disruptive Disorders		
ADHD	8	4,8%
ODD	6	3,6%
CD	0	0%
Mood Disorders		
Current Depressive Episode	9	5,4%
Past Depressive Episode	8	4,8%
Current (Hypo)Manic Episode	0	0%
Past (Hypo)Manic Episode	0	0%
Current Dysthymic Disorder	4	2,4%
Depressive Disorder[1]	15	9%
Depressive Disorder NOS	3	1,8%
Current Depressive Symptoms1	14	8,4%
Anxiety Disorders		

Table 2

Frequencies of diagnoses based on the KID-SCID interview (N = 167) (continued)

Diagnosis	N	Percentage
Separation Anxiety Disorder	0	0%
Social Phobia	0	0%
Specific Phobia	2	1,2%
OCD	0	0%
PTSD	2	1,3%
GAD	1	0,6%
Panic Disorder	0	0%
Agoraphobia	0	0%

ADHD = Attention Deficit Hyperactivity Disorder, ODD = Oppositional Defiant Disorder, CD = Conduct Disorder, OCD = Obsessive-Compulsive Disorder, PTSD = Post Traumatic Stress Disorder, GAD = General Anxiety Disorder

[1] Note: In the current study we used Current Depressive Symptoms as dependent variable, including all youngsters who were diagnosed with a Major Depressive Disorder with a current episode, a Dysthymic Disorder with current symptoms or Depressive Disorder NOS with current symptoms. Since the Depressive Disorder affects also youngsters who experienced a depressive episode in the past we could not use this diagnostic category in our analyses.

First, to test whether two measure moments better explain the variance in depressive symptoms a binary logistic regression was conducted with Current Depressive Symptoms (CDS) as a categorical dependent variable. In a first block we included the total score on the second measure moment of the CDI (CDI-T2). Since the 'CDS' and 'no-CDS' group differed in gender rate, we also included gender as a predictor variable in the first block. The total score of the first measure moment of the CDI (CDI-T1) was added as a predictor in the second block.

To control for comorbid psychopathology, we conducted a regression analysis to predict CDI-T2-scores based on all the RCADS anxiety subscales and YSR-externalising behaviour problems scale, and saved the standardised scores as a new variable. We repeated the first regression analysis, now with CDS as dependent variable, gender, the standardised comorbidity score and CDI-T2 as predictors in the first block and the CDI-T1 score added as predictor in the second block.

We choose to add the CDI-T1 at the last block, because this measurement was most separated in time from the clinical interview. However as a check, we also ran the analyses with the CDI-T1 as predictor in the first block, and CDI-T2 in the second block.

Second, to test the second research question, we conducted a binary logistic regression with CDS as a categorical dependent variable, with gender as predictor in the first block and the addition of the parent reports for internalising problems (CBCL-INT) in the second block. To test whether parent

reports had an addition predictive value on top of child reports on depressive symptoms we conducted a logistic regression analysis with CDI-T2 and gender as predictors in the first block and CBCL-INT added in the second block.

To control for comorbid psychopathology, we repeated the regression analysis with CDS as dependent variable, gender and the standardised comorbidity score and CDI-T2 score as predictors in the first block and added the CBCL-INT in the second model.

Again we repeated all analyses of the second research question with CDI-T1 as a predictor in the models instead of CDI-T2 to control whether different measure moments changed the results of the logistic regression analyses.

Results

Descriptives

In this study, 98% of the youngsters and 76% of the parents completed all the questionnaires. With regard to the demographic variables we found no differences between the 'CDS' (n = 14) and 'no-CDS' group (n = 153) for mean age ($F(1,163) = .33$, $p = .57$), school level ($\chi^2(3) = .61$, $p = .89$), and SES ($\chi^2(3) = 3.43$, $p = .33$). We did find a relationship between CDS and gender (see Table 3), with female gender being more present in the CDS-group.

Table 3
Descriptive Statistics

	CDS	No-CDS	Statistics
Gender	n_{boys} = 2 (14%) n_{girls} = 12 (86%)	n_{boys} = 71 (46%) n_{girls} = 82 (54%)	$\chi^2(1) = .5.38$, $p < .05$
CDI-T1	M = 22.97, SD = 6.07	M =11.97, SD = 7.81	$F(1,165) = 24.47$, $p < .001$
CDI-T2	M = 21.68, SD = 6.98	M = 9.19, SD = 6.61	$F(1,165) = 45.38$, $p < .001$
YSR internalising	M = 30.86, SD = 8.90	M =13.63, SD = 9.58	$F(1,165) = 41.91$, $p < .001$
YSR externalising	M = 17.57, SD = 8.09	M =9.92, SD = 7.69	$F(1,165) = 12.59$, $p < .01$
CBCL internalising	M = 11.42, SD = 8.20	M =5.40, SD = 4.50	$F(1,165) = 19.41$, $p < .001$
CBCL externalising	M = 9.97, SD = 5.60	M =5.61, SD = 5.78	$F(1,165) = 7.33$, $p < .01$
RCADS SP	M = 7.58, SD = 5.43	M =2.87, SD = 3.60	$F(1,165) = 19.92$, $p < .001$
RCADS PD	M = 4.73, SD = 3.07	M =1.76, SD = 2.67	$F(1,165) = 15.50$, $p < .001$
RCADS SAD	M = 3.41, SD = 3.90	M =1.15, SD = 2.06	$F(1,165) = 12.84$, $p < .001$
RCADS GAD	M = 4.77, SD = 3.38	M =1.82, SD = 2.75	$F(1,165) = 14.17$, $p < .001$
RCADS OCD	M = 3.19, SD = 2.30	M =1.40, SD = 2.43	$F(1,165) = 7.04$, $p < .01$

Note: CDI-T1 = Children's Depression Inventory Measure Time 1, CDI-T2 = Children's Depression Inventory Measure Time 2, SP = Social phobia, PD = panic disorder, SAD = separation anxiety disorder, GAD = generalised anxiety disorder, OCD = obsessive-compulsive disorder

Furthermore we compared the 'CDS' and 'no-CDS' group on the psychopathology self-report questionnaires (see Table 3). As expected, both groups differed on both the CDI-T1 and CDI-T2 total scores. We observed higher CDI scores in the CDS-group. We further found significant differences for the child reports on YSR-internalising and YSR-externalising problems and for the parent reports on CBCL-internalising and CBCL-externalising problems; with more emotional and behavioural problems in the CDS-group. Finally, on the RCADS we found differences between both groups for all the subscales with more symptoms of anxiety symptoms in the CDS-group including scales of "Social phobia", panic disorder, separation anxiety disorder, generalised anxiety disorder, and obsessive-compulsive disorder.

1. Can we make a better prediction of diagnoses in variance of depressive symptoms based on two assessment moments?

The first model, with gender and CDI-T2 as predictors, significantly explained the presence of CDS ($\chi^2(2) = 37.84$, $p < .001$) and accounted for between 20.5% and 48.3% of the variance in CDS, with 98.7% of the non-depressed and 30.8% of the depressed youngsters successfully predicted. Overall 93.3% of the predictions were correct. The β-coefficients associated with the CDI-T2 total score ($\beta = .25$, $Wald(1) = 17.49$, $p < .001$) and gender ($\beta = -1.82$, $Wald(1) = 4.19$, $p < .05$) were significantly different from zero. The odds ratio ($Exp\ \beta$) of 1.28 for the CDI-T2 indicated that there is a 1.28-point increase in risk for Current Depressive Symptoms associated with each 1-point increase in the CDI score. The odds ratio for gender was somewhat lower ($Exp\ \beta = .16$). The second model, with CDI-T1 added as predictor, also predicted significantly the presence of CDS ($\chi^2(3) = 38.30$, $p < 0.001$). The model accounted for between 20.7% and 48.9% of the variance in CDS, and 98.7% of predictions for the non-depressed and 30.8% for the depressed group was accurate. Overall percentage of correct predictions for the second model was 93.3%. The β-coefficients associated with the CDI-T2 total score and gender were still significantly different from zero, $\beta = .23$, $Wald(1) = 11.82$, $p = .001$ for CDI-T2 and $\beta = -1.76$, $Wald(1) = 3.89$, $p < .05$ for gender whereas the β-coefficient associated with the CDI-T1 total score did not reach the significance level ($Wald(1) = .52$, $p = .47$). $Exp\ \beta$ was 1.26 for CDI-T2 and .17 for gender. Although both models are significant in predicting the presence of CDS, the addition of a second assessment moment of the CDI (CDI-T1) in the second model does not cause any increase in the percentage of correct predictions. The same pattern of results was found when repeating the regression with CDI-T2 replaced by CDI-T1 and vice versa.

As regards the regression with the comorbid psychopathology included (see Table 4 and 5), both models were significant in explaining CDS with

93.3% as overall correct predictions. In line with the first analysis, for both the first model and second model only the β-coefficients associated with the CDI-T2 total score and gender were significantly different from zero. The new standardised comorbidity score did not explain the presence of CDS significantly. Again, the same pattern of results was found when repeating the regression with CDI-T2 replaced by CDI-T1 and vice versa.

Table 4

Logistic regression analysis with CDS as dependent variable

	B	**Wald**	***df***	***p***	**Exp(*β*)**	***χ²***
Block 1						**38.31 (3)*** **
Gender	-1.90	4.33	1	.04*	.15	
Comorbidity	.22	.47	1	.49	1.25	
CDI-T2	.22	9.07	1	.00**	1.25	
Block 2						**38.53 (4)*** **
Gender	-1.83	3.98	1	.04*	.16	
Comorbidity	.17	.24	1	.63	1.18	
CDI-T2	.21	8.12	1	.00**	1.24	
CDI-T1	.03	.24	1	.62	1.03	

Note: Comorbidity = new standardised scores for comorbidity; CDI = Children's Depression Inventory
*=p <.05; **=p <.01; *** = p <.001

Table 5

Cross Table with the actual and predicted CDS-rates based on gender, comorbidity and CDI-T1&T2

		Predicted	
Observed	No-CDS	CDS	**Percentage correct**
No-CDS	151	2	98.7
CDS	9	5	35.7
Overall percentage			93.4

2. Can we make a better prediction of diagnoses in variance of depressive symptoms based on a multi-informant method?

The first model, with gender as the only predictor, ($\chi^2(1) = 6.07$, $p < .05$) as well as the second model, with CBCL-INT added ($\chi^2(2) = 19.21$, $p < .001$) significantly explained the presence of CDS. The second model accounted for between 10.9% and 24.9% of the variance in CDS, with 99.3% of the non-depressed and 14.3% of the depressed youngsters successfully predicted. Overall 92.9% of the predictions were correct. Both gender ($\beta = -1.77$, $Wald(1) = 4.74$, $p < .05$) and the CBCL-INT ($\beta = .17$, $Wald(1) = 10.90$, $p <$

.001) had a β-coefficient significantly different from zero with an odds ratio (Exp β) of .17 for gender and 1.18 for CBCL-INT.

To test whether parent reports added to the predictive value of child reports of depressive symptoms we conducted a logistic regression analysis with CDI-T2 and gender as predictors in the first block and CBCL-INT added in the second block. In this second analysis, the first model significantly predicted the presence of CDS ($\chi^2(2) = 39.83$, $p < .001$). The model accounted for between 21.2% and 48.5% of the variance in CDS. Overall 92.8% of the predictions were correct. Both gender ($\beta = -1.94$, $Wald(1) = 4.85$, $p < .05$) and the CDI-T2 total score ($\beta = .25$, $Wald(1) = 18.12$, $p < .001$) had a β-coefficient significantly different from zero with an *Exp β* of .14 for gender and 1.28 for CDI-T2 total score. The second model also predicted significantly the presence of CDS ($\chi^2(3) = 42.50$, $p < 0.001$). The model accounted for between 22.5% and 51.3% of the variance in CDS, with 98.7% accurate predictions for the non-depressed and 42.9% for the depressed group. Overall percentage of correct predictions for the second model was 94%. The β-coefficient associated with gender, $\beta = -2.13$, $Wald(1) = 4.92$, $p < .05$ and the CDI-T2 total score was still significantly different from zero, $\beta = .22$, $Wald(1) = 14.32$, $p < .001$. Whereas the β-coefficient associated with CBCL-internalising problems did not reach the significance level. *Exp β* for gender was .12 and 1.25 for CDI-T2. Both models are significant predictors of predicting CDS and the addition of CBCL-INT results in a higher overall percentage of correct predictions. However, the parent report on internalising problems as such is not a significant predictor of CDS on top of the child reports on depressive symptoms (CDI-T2).

As regards the regression with the comorbid psychopathology included (see Table 6 and 7), both models were significant in predicting CDS. The percentage of overall correct predictions of the first model (93.4%) was somewhat lower than the one for the second model (94.0%). In contrast to the β-coefficient associated with the gender ($\beta = -2.14$, $Wald(1) = 4.94$, $p < .05$) and CDI-T2 total score ($\beta = .21$, $Wald(1) = 9.32$, $p < .01$), the coefficient associated with the CBCL-INT did not differ significantly from zero.

Also for the second research question, to control whether the order of the measure moments changed the results of the logistic regression analyses, we repeated all the analyses with CDI-T1 replacing the CDI-T2 as a predictor, with no significant differences in the results.

Table 6
Logistic regression analysis with CDS as dependent variable

	B	Wald	df	p	Exp(β)	χ^2
Block 1						**40.22 (3)*** **
Gender	-2.02	4.97	1	.03*	.13	
Comorbidity	.20	.40	1	.53	1.22	
CDI-T2 total score	.22	9.61	1	.00**	1.25	
Block 2						**42.58 (4)*** **
Gender	-2.14	4.94	1	.03*	.12	
Comorbidity	.09	.09	1	.77	1.10	
CDI-T2 total score	.21	9.32	1	.00**	1.24	
CBCL Intern. Prob.	.09	2.29	1	.13	1.09	

Comorbidity = new standardised scores for comorbidity; CDI = Children's Depression Inventory; CBCL = Children's Behaviour Check List
* = $p < .05$; ** = $p < .01$; *** = $p < .001$

Table 7
Cross Table with the actual and predicted CDS-rates based on gender, comorbidity, CDI-T2 and CBCL internalising problems

		Predicted	
Observed	No-CDS	CDS	**Percentage correct**
No-CDS	151	2	98.7
CDS	8	6	42.9
Overall percentage			94.0

Discussion

The current study extends the existing literature on early identification of depression in children and adolescents. We evaluated the accuracy of a one-wave versus two-wave assessment as well as the value of adding multi-informant data in a non-clinical sample of 167 youngsters between 10 and 15 years old for screening depressive symptoms. A structured clinical interview was administered in a large non-clinical sample with a broad age range, and this was judged twice by independent encoders to ensure the reliability of the diagnoses.

First, in contrast to our expectations, the logistic regression analyses showed that two measure moments of the CDI are not more accurate in capturing depressive symptoms compared to one. Even when controlling for gender and comorbid anxiety and externalising behaviour problems one measure moment of the CDI is equally accurate in predicting a diagnosis of Current Depressive Symptoms as a two wave measure. The results of the current study are inconsistent with Lewinsohn and Teri (1982) and Roberts et al.

(1991) who found improved discriminant validity after the serial administration of the same rating scale. Perhaps the differences in results are due to methodological differences. Roberts and his colleagues used another rating scale, namely the Center for Epidemiologic Studies-Depression Scale. More important, both studies used a larger time interval from a mean of 9 days in Roberts et al. (1991) and 3 to 6 months in Lewinsohn et al (1982). Roberts even found a discernable effect on discriminant validity among subjects with intervals until 31 days.

Although we could not prove a statistically better capture of depressive symptoms based on two measure moments, it is still possible that a second administration of the CDI is valuable in the clinical practice. As already stated before there can be an exaggeration or an underreporting of the depressive symptoms at the first administration compared to the clinicians judgement. In a second administration children have already filled in the questionnaire, understand better the questions, had the time to think about what the items really mean in their lives,… Future investigations into the benefits of serial use of the CDI in a multistage strategy for depression in children is necessary. A larger time interval could result in a confirmation of our hypothesis that two measure moments improve the accuracy in detecting depressive symptoms. However, it was our intention to examine daily fluctuations in depression scores for which an interval of one week is adequate.

Second, the overall correct prediction of depressive symptoms with percentages between 92.9% and 94% was excellent. However this percentage mainly derives from very good prediction of the non-depressed youngsters. Only 30.8 to 42.9% of the depressed youngsters were correctly identified which results in a low specificity. This low specificity is in contrast with the specificity rate of 83.8% of Timbremont, Braet en Dreesen (2004) who investigated the utility of the CDI for predicting a diagnosis of a depressive disorder based on the KID-SCID. Conversely their sample included older adolescents until 18 years and they selected referred youngsters. Since we wanted to evaluate a screening procedure for depressive symptoms in children who don't call assistance by themselves, it would not have been an appropriate method to use a clinical sample. From these results, we can conclude that in younger children provisionally without clearly visible symptoms, clinicians should not only rely on one measure moment of a self report screening instrument and a clinical interview can be an interesting supplementation. Since clinical interviews are too time consuming, further investigation is needed to make decision criteria for administrating a clinical interview among at risk youngsters. Although the CDI is adapted in language and catches the core symptoms as well as all the symptoms that have a differential expression as a function of age and development, it cannot fully identify potential cases to be interviewed for diagnostic purposes in younger school samples.

Third, we wanted to explore the additional value of parent reports in explaining variance in the presence of Current Depressive Symptoms among their children. Based on the existing literature, child reports are recognised as important information sources, but could be insufficient among children of young age (Rubio-Stipec, Fitzmaurice, Murphy & Walker, 2003). In line with previous research, it was assumed that parent reports on emotional and behavioural problems account for a better and unique prediction of variance in depressive symptoms (Ferdinand et al., 2003; Verhulst, et al., 1997). Parents are more involved in the day-to-day lives of children and therefore know more about their behaviour and activities (e.g. compared to teachers). Regarding the second aim of the study, the results showed however that there is no clear evidence for or against using a multi-informant method in the screening for depressive symptoms among children and youngsters. Although a model with parent reports on internalising problems was statistically significant in predicting depressive symptoms, the parent reports identified only 14.3% of the depressed youngsters (n = 2) to get a diagnosis of Current Depressive Symptoms. This is a very low rate of correct identified positive cases. In addition, the parent reports themselves were no significant predictors of variance in depressive symptoms on top of child reports of depressive symptoms measured by the CDI.

Unfortunately, the current study could not replicate the results of Ferdinand and his colleagues (2003) and Verhulst and colleagues (1997) who reported a unique contribution of parent reports on top of child reports in predicting child's psychopathology. How can we explain these unexpected findings? First, the parent report questionnaires used in the present research have strong psychometric qualities (Verhulst & Van der Ende, 2004). So, we assume that we cannot explain our falsified hypothesis by questioning the validity of the parent measure. Second, we used a broad-band parent rating of internalising problems, instead of a specific screening instrument for depressive symptoms. It would be interesting to determine whether a parental measure that specifically assesses depression instead of a broad band questionnaire would improve the predictive ability. Research on the English version of the parent version of the CDI (CDI-P) resulted in a high degree of internal consistency, good test-retest reliability, and a moderately high degree of convergent validity (Cole, Truglio, & Peeke, 1997; Kazdin, French, Unis, & Esveldt-Dawson, 1983; Wierzbicki, 1987). In future research it might be interesting to further investigate the role of parent reports. For example, the current study could be replicated and include the CDI-P in the test-battery. We must acknowledge that in this study both parent and child reports have almost equal good rate of predicting variance in depressive symptoms.

In the current study gender appeared to be the most important predictor of depressive symptoms and accounted for a big amount of variance. However

this information is not very valuable for clinicians to use in screening procedures. Therefore all analyses were repeated without the inclusion of gender but did not reflect in other conclusions about the additional value of a second measurement moment neither for adding parent reports.

This study has some limitations to mention. First, we could question the validity of the outcome measure. The KID-SCID is an investigator-based measure, so the interviewer is responsible for rating the criteria, and this was based on child responses only. Overall the first results indicate the KID-SCID as a good clinical instrument with substantial to outstanding agreement between different raters. Though the Cohen's Kappa for "past depressive episode" was remarkably lower (.32) compared to the other disorders. Previous research already showed that younger children have difficulties reporting on information requiring temporal parameters; and other informants must be relied on for information on course such as age of onset, previous episodes and duration of current episode (Kovacs, 1986). It could be possible that children find it even more difficult to describe these factors when they need to report retrospectively over a long period of time, and therefore causes disagreement between interviewers. Therefore the diagnoses could have differed when parent information was also at the examiner's disposal. Moreover, since the diagnoses were only based on the child's information and it was not amalgam of child and parent reports, it is logical that we see higher agreements between the child self-report measures of depressive symptoms (CDI) and the clinical interview administered with the children. Possibly when we had used both parent and child information in diagnosing the children, parent reports could have had a better predictive value in explaining the presence of Current Depressive Symptoms.

Second, we can describe a few weaknesses with regard to the method. Although this was not a goal of the current study, the small interval between different measure moments of this prospective design makes it impossible to investigate the change in symptoms over time. Parent-child agreements of symptom change over time appeared to be considerably better than agreements about the level of children's depression at a specific point in time (Cole, et al., 2002). Future longitudinal research to the predictions in time or the surplus value of parent reports for symptom change reports is recommended. Next, the non-clinical character of this study sample caused low rates of diagnoses, this could affect the predictive ability of the CDI. Further, due to limited time and cost we could not administer the KID-SCID in all the subjects of the study. However, we did choose a very low cut-off score (CDI total score 10) to decide who should be interviewed. We originally planned to interview an equal number of pupils below as above the cut-off. Unfortunately this was not possible due to limited time space that was provided by the schools. We interviewed 64 children under the cut-off to control for false-

negatives. None of these 64 showed symptoms of depression or other mood disorders. We also investigated whether these 64 participants were representative for the rest of the "below cut-off group" (n = 244). We found no differences for age, school type, and SES. We did found a difference for gender (χ^2 (244) = 6,91, $p < .01$), with more girls in the interviewed group compared to the group that was not interviewed.

In addition, we did not include teacher rating scales although these informants also could have served important unique information about the child's or adolescent's problems.

In summary, we can conclude that one measure moment is equally reliable for catching the core symptoms of depression and to give an early indication of possible depression among children and youngsters than two measure moments. We could identify14 children as suffering from severe depressive symptoms that met all the criteria for a depressive disorder. This means that these children reported symptoms of depression as well as impaired functioning during the KID-SCID interview. Yet, they or their parents did not call assistance by themselves, demonstrating the relevance of early identification. Unfortunately, a self-report questionnaire could not identify all of the depressed cases. The current results also demonstrate that parents' reports are no clear predictors of their children's depressive symptoms. Therefore future research into screening procedures stays a priority for clinical practice. In younger children we cannot only rely on self-report measurements, and the addition of a clinical interview is valuable. Though interviews are too time-consuming, further research is necessary to find clear markers for at risk children.

References

Achenbach, T.M., & Rescorla, L.A. (2001). Manual for the ASEBA School-Age Forms & Profiles. Burlington, VT: University of Vermont, Research Center for Children, Youth, & Families.

Beck, A.T., Ward, C.H., Mendelson, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry, 4,* 561-571.

Birmaher, B., Brent, D., & AACAP Workgroup on Quality Issues (2007). Practice parameter for the assessment and treatment of children and adolescents with depressive disorders. *Journal of the American Academy of Child and Adolescents Psychiatry, 46,* 1503-1526.

Braaten, E.B., Biederman, J., DiMauro, A., Mick, E., Monuteaux, M.C., Muehl, K., & Faraone, S.V. (2001). Methodological complexities in the diagnosis of major depression in youth: an analysis of mother and youth self-reports. *Journal of Child and Adolescence Psychopharmacology, 11*, 395-407.

Chorpita, B.F., Yim, L., Moffitt, C.E., Umemoto, L.A., & Francis, S.E. (2000). Assessment of symptoms of DSM-IV anxiety and depression in children: A

revised child anxiety and depression scale, *Behaviour Research and Therapy, 38*, 835-855.

Cole, D.A., Truglio, R., & Peeke, G.L. (1997). Relation between symptoms of anxiety and depression in children: A multitrait-multimethod-multigroup assessment. *Journal of Consulting and Clinical Psychology, 65,* 110-119.

Cole, D.A., Tram, J.M., Martin, J.M, Hoffman, K.B., Ruiz, M.D., Jacquez, F.M., & Maschman, T.L. (2002). Individual differences in the emergence of depressive symptoms in children and adolescents: A longitudinal investigation of parent and child reports. *Journal of Abnormal Psychology*, *111,* 156-165.

Connell, A.M., & Goodman, S.H. (2002). The association between psychopathology in fathers versus mothers and children's internalising and externalising behavior problems: A meta-analysis. *Psychological Bulletin, 128*, 746-773.

Craighead, W.E., Smucker, M.R., Craighead, L.W., & Ilardi, S.S. (1998). Factor analysis of the Children's Depression Inventory in a community sample. *Psychological Assessment, 10*(2), 156-165.

De Los Reyes, A., Goodman, K.L., Kliewer, W., & Reid-Quinones, K. (2008). Whose depression relates to discrepancies? Testing relations between informant characteristics and informant discrepancies from both informants' perspectives. *Psychological Assessment, 20,* 139-149.

De Los Reyes, A., & Kazdin, A.E. (2005). Informant discrepancies in the assessment of childhood psychopathology: A critical review, theoretical framework, and recommendations for further study. *Psychological Bulletin, 131,* 483-509.

Dreessen, L., Stroux, A., & Weckx, M. (1998). *Nederlandse vertaling van het Gestructureerd Klinisch Interview voor DSM-IV-kind versie (KID-SCID; Versie 1.0)* [Dutch translation of the Structured Clinical Interview for *DSM-IV* Child edition]. Maastricht, The Netherlands: Author.

Fendrich, M., Weissman, M.M., & Warner, V. (1990). Screening for Depressive Disorder in Children and Adolescents: Validating the Center for Epidemiologic Studies Depression Scale for Children. *American Journal for Epidemiology, 131,* 538-551.

Ferdinand, R.F, Hoogerheide, K.N., van der Ende, J., Visser, J.H., Koot, H.M., Kasius, M.C., & Verhulst, R.C. (2003). The role of the clinician: Three-year predictive value of parents' teachers', and clinicians' judgments of childhood psychopathology. *Journal of Child Psychology and Psychiatry, 44,* 867-876.

Garber, J., & Kaminski, K.M. (2000). Laboratory and performance-based measures of depression in children and adolescents. *Journal of Clinical Child Psychology, 29,* 509-525.

Garrison, C.Z, Addy, C.L., Jackson, K.L., McKeown, R.E., & Waller, J.L. (1991). The CES-D as a screen for depression and other psychiatric disorders in adolescents. *Journal of American Academy of Child and Adolescent Psychiatry, 30*(4), 636-641.

Garrison, C.Z., Jackson, K.L., Marsteller, F., McKeown, R., & Addy, C. (1990). A longitudinal study of depressive symptomatology in young adolescents. *Journal of American Academy of Child and Adolescent Psychiatry, 29*(4), 581-585

Goodman, R., Ford, T., Simmons, H., Gatward, R., & Meltzer, H. (2003). Using the Strengths and Difficulties Questionnaire (SDQ) to screen for child psychiatric

disorders in a community sample. *International Review of Psychiatry. 15*(1-2), 166-172.

Harrington, R., Rutter, M., & Fombonne, E. (1996). Developmental pathways in depression: Multiple meanings, and endpoints. *Development and psychopathology, 8,* 606-616.

Hien, D., Matzner, E.J., Eirst, M.B., Spitzer, R.L., Gibbon, M., & Williams, J.B.W. (1994). *Structured Clinical Interview for DSM-lV-Child edition (Version 1.0).* New York: Columbia University.

Hollingshead, A. (1975). *Four factor index of social status.* New Haven, CT: Yale University Press.

Jensen, R.S., Rubio-Stipec, M., Canino, G., Bird, H.R., Dulcan, M.K., Schwab-Stone, M.E., & Lahey, B.B. (1999). Parent and child contributions to diagnosis of mental disorder: Are both informants always necessary? *Journal of the American Academy of Child & Adolescent Psychiatry, 38,* 1569-1579.

Kazdin, A.E., French, N.H., Unis, A.S., & Esveldt-Dawson, K. (1983). Assessment of childhood depression: Correspondence of child and parent ratings. *Journal of the American Academy of Child Psychology, 22,* 157-164.

Kendall, P.C., Cantwell, D.P., & Kazdin, A.E. (1989). Depression in children and adolescents: Assessment issues and recommendations. *Cognitive Therapy and Research, 13,* 109-146.

Klein, D.N., Dougherty, L.R., & Olino, T.M. (2005). Towards guidelines for evidence-based assessment of depression in children and adolescents. *Journal of Clinical Child and Adolescence Psychology, 34,* 412-432.

Kovacs, M, (1986), A developmental perspective on the methods and measures in the assessment of depressive disorders: The clinical interview. In M. Rutter, C.-E. Izard, & P.B. Read (Eds.), *Depression in young people: Developmental and clinical perspectives* (pp. 435-465), New York: Guilford,

Kovacs, M. (1992). *Children's Depression Inventory (CDI).* New York: Multi-Health Systems, Inc.

Lagges, A.M., & Dunn, D.W. (2003). Depression in children and adolescents. *Neurologic Clinics of North America, 21,* 953-960.

Landis, J.R., & Koch, G.G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33,* 159-174.

Lewinsohn, P.M., & Teri, L. (1982). The selection of depressed and nondepressed subjects on the basis of self-report data. *Journal of Consulting and Clinical Psychology, 50,* 590-591.

Little, R. (1988). A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association, 83,* 1198-1202.

Matzner, E. (1994). *Videotapes as training tools for the development of the KID-SCID.* Paper presented at the scientific proceedings of the American Academy of Child and Adolescent Psychiatry, 42nd annual meeting. Washington DC.

Muris, P., Meesters, C., & Schouten, E. (2002). A brief questionnaire of DSM-IV defined anxiety and depression symptoms among children. *Clinical Psychology and Psychotherapy, 9,* 430-442.

Muris, P., Schmidt, H., Merckelbach, H., & Schouten, E. (2001). The structure of negative emotions in adolescents. *Journal of Abnormal Child Psychology*, *29*, 331-337.

Myers, K., & Winters, N.C. (2002). Ten-Year Review of Rating Scales II: Scales for Internalising Disorders. *Journal of American Academy of Child and Adolescent Psychiatry, 41,* 634-659.

Nelson, W.M., & Politano, P.M. (1990). Children's Depression Inventory: Stability over repeated administration in psychiatric inpatient children. *Journal of Clinical Child & Adolescent Psychology, 19,* 254-256.

O'Brien, R.M. (2007). A caution regarding rules of tumb for Variance Inflation Factors. *Quality and quantity, 41,* 673-690.

Ollendick, T.H., Grills, A.E., & King, N.J. (2001). Applying developmental theory to the assessment and treatment of childhood disorders: Does it make a difference? *Clinical Psychology & Psychotherapy, 8*(5)*,* 304-314.

Puura, K., Almqvist, F., Tamminen, T., Piha, J., Kumpulainen, K., Rasanen, E., Moilanen, I., & Koivisto, A.M. (1998). Children with symptoms of depression – What do the adults see? *Journal of Child Psychology and Psychiatry, 39,* 577-585.

Reynolds, W.M. (1986). A model for the screening and identification of depressed children and adolescents in school settings. *Professional School Psychology, 1,* 117-129.

Roberts, R.E, Lewinsohn, P.M., & Seeley, J.R. (1991). Screening for adolescent depression: A comparison of depression scales. *Journal of American Academy of Child and Adolescent Psychiatry, 30,* 58-66.

Roelofs, J., Braet, C., Rood, L., Timbremont, B., van Vlierberghe, L., Goossens, L., & van Breukelen, G. (2010). Norms and screening utility of the Dutch version of the children's depression inventory in clinical and nonclinical youth. *Psychological Assessment, 22,* 866-877.

Rubio-Stipec, M., Fitzmaurice, G., Murphy, J., & Walker, A. (2003). The use of multiple informants in identifying the risk factors of depressive and disruptive disorders – Are they interchangeable? *Social Psychiatry and Psychiatric Epidemiology, 38*, 51-58.

Schafer, J. (1997). *Analysis of incomplete multivariate data.* London: Shapman and Hall.

Spence, S.H. (1998). A measure of anxiety symptoms among children. *Behavior Research and Therapy, 36,* 545-566.

Spitzer, R.L., Williams, J.B.W., & Gibbon, M. (1986). *Structured clinical interview for* DSM-III-R. New York: New York State Psychiatric Institute, Biometrics Research.

Spitzer, R.L., Williams, J.B.W., Gibbon, M., & First, M. (1992). The structured clinical interview for *DSM-III-R* (SCID) I: History, rationale, and description. *Archives of General Psychiatry, 49,* 624-629.

Stark, K.D. (1990). *Childhood depression: school-based intervention (*pp. 68)*.* New York: Guilford Press.

Stark, K.D., Humphrey, L.L., Laurent, J., Livingston, R., & Christopher, J. (1993). Cognitive, behavioural, and family factors in the differentiation of depressive

and anxiety disorders during childhood. *Journal of Consulting and Clinical Psychology, 61,* 878-886.

Timbremont, B., & Braet, C. (2002). *Children's Depression Inventory: Nederlandstalige versie* [Children's Depression Inventory: Dutch version]. Lisse: Swets & Zeitlinger.

Timbremont, B., Braet, C., & Dreesen, 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,* 149-157.

US Preventive Services Task Force (2002). Screening for depression: recommendations and rationale. *Annual Internal Medicine, 136,* 760-764.

Verhulst, F.C., Dekker, M.C., & van der Ende, J. (1997). Parent, teacher, and self-reports as predictors of signs of disturbance in adolescents: Whose information carries the most weight? *Acta Psychiatrica Scandinavica, 96,* 75-81.

Verhulst, F.C., & Van der Ende, J. (2004). *Handleiding voor de CBCL/6-18 and YSR [manual for the child behavior checklist and revised child behavior profile and youth self report and profile]*. Sophia Children's Hospital, Department of Child and Adolescent Psychiatry, Rotterdam.

Wierzbicki, M. (1987). A parent form of the Children's Depression Inventory: Reliability and validity in nonclinical populations. *Journal of Clinical Psychology*, *43,* 390-397.

Williams, J.B., Gibbon, M., First, M.B., & Spitzer, R.L. (1992). The structured clinical interview for *DSM-III-R* (SCID): II. Multisite test-retest reliability. *Archives of General Psychiatry, 49,* 630-636.

Williams, S.B., O'Connor, E.A., Eder, M., & Whitlock, E.P. (2009). Screening for child and adolescent depression in primary care settings: a systematic evidence review for the U.S. Preventive Services Task Force. *Pediatrics, 123,* 716-723.

Yorbik, O., Birmaher, B., Axelson, D., Williamson, D.E., & Ryan, N.D. (2004). Clinical characteristics of depressive symptoms in children and adolescents with major depressive disorder. *Journal of Clinical Psychiatry, 65*, 1654-1659.

Zeman, J., Klimes-Dougan, B., Cassano, M., & Adrian, M. (2007). Measurement Issues in Emotion Research With Children and Adolescents. *Clinical Psychology: Science and Practice, 14*(4), 377-401.

Zwaanswijk, M., Verhaak, P.F., Bensing, J.M., van der Ende, J., & Verhulst, F.C. (2003). Help seeking for emotional and behavioural problems in children and adolescents: a review of recent literature. *European Child and Adolescent Psychiatry, 12,* 153-161.

Received: August 14, 2012
Revision received: February 14, 2013
Accepted: March 12, 2013