

**ASSESSING EMOTIONAL AND BEHAVIOURAL  
PROBLEMS WITH THE CHILD BEHAVIOUR CHECKLIST:  
EXPLORING THE RELEVANCE OF ADJUSTING THE  
NORMS FOR THE FLEMISH COMMUNITY**

Caroline BRAET, Justine CALLENS, Mark SCHITTEKATTE,  
Veerle SOYEZ, Celina DRUART & Herbert ROEYERS\*  
*Ghent University*

The Child Behaviour Checklist (CBCL) is one of the most frequently used dimensional instruments for screening behavioural and emotional problems in children. In this study the psychometric properties of the CBCL 2001-version and the usefulness of existing US norms within a Flemish community sample were explored. Mothers of young children ( $N = 170$ ) and school-aged children and adolescents ( $N = 718$ ) completed the CBCL and the Strengths and Difficulties Questionnaire (SDQ). Reliability of both the CBCL/1½-5 and CBCL/6-18 was excellent. There were substantial correlations between the SDQ and CBCL scales, indicating a good concurrent validity. Significant differences were found when comparing the mean CBCL raw scores of the different subsamples (for different ages and gender groups) with the US norms. Generally spoken, the mean CBCL raw scores as well as the clinical and borderline clinical cutpoints for Flemish children were mostly lower, specifically for the Total Problems score and the Externalising scales.

**Introduction**

It is generally accepted that assessment of psychological problems in children and adolescents should not be delayed until these problems reach more serious levels and more intensive and expensive interventions are required (Braet & van Aken, 2006, Costello & Angold, 2000; Culbertson, 1993; Tick, van der Ende, & Verhulst, 2007). Additionally, the high prevalence rates of psychopathology among children as demonstrated by research in random samples, compared to the relatively low number of referrals to specialised care, illustrate that it is crucial to screen children in community samples (Burns, Philips, Wagner, Barth, Kolko, Campbell et al., 2004). To identify at-risk

\* Caroline Braet, Department of Developmental, Personality, and Social Psychology; Justine Callens and Mark Schittekatte, Testpracticum; Veerle Soyez, Department of Special Education; Celina Druart, Department of Data Analysis; Herbert Roeyers, Department of Experimental Clinical and Health Psychology. All authors are affiliated with Ghent University.

We thank the Faculty of Psychology and Educational Sciences of Ghent University for their financial support.

Correspondence concerning this article should be addressed to Caroline Braet, Department of Developmental, Personality, and Social Psychology; H. Dunantlaan 2, B-9000 Gent. E-mail: Caroline.Braet@UGent.be

children, mental health professionals need reliable and valid screening methods for behavioural and emotional problems in children. However, screening through observation or interview is a great burden and these methods are generally neither very reliable nor cost-effective (Jensen & Weisz, 2002). Therefore, rating scales with clear cutpoints are seen as 'golden standards' for monitoring youth with mental problems and to select potential cases for further assessment. However, the appropriateness of these instruments is still under study.

Determining cutpoints for screening instruments implies they have been tested on their sensitivity and specificity to identify cases, whereby the number of false positives and false negatives is kept as low as possible. Moreover, the norm group upon which they rely must meet the highest criteria regarding representativeness according to the community they refer to. Although for the Flemish community (Dutch speaking part of Belgium) norm groups are already available for some small-band problems like disruptive behaviour disorders (Oosterlaan, Baeyens, Scheres, Antrop, Roeyers, & Sergeant, 2008) or depression (Timbremont, Braet, & Roelofs, 2008), hardly any norm groups are available for broad-band screening instruments in Dutch speaking samples. Up to now, only one study (Van Leeuwen, Meerschaert, Bosmans, De Medts, & Braet, 2006) discussed the use of a general screening instrument and provided norms for 4- to 7-year-old children. In the meantime, screening instruments developed in US are widely used in clinical practice without investigating empirically whether the norms and cutpoints of these instruments are appropriate for the Flemish community. Although some impressive studies already tested the robustness of screening instruments across diverse cultures (Achenbach, Becker, Döpfner, Heiervang, Roessner, Steinhausen et al., 2008; Crijnen, Achenbach, & Verhulst, 1997; Ivanova, Achenbach, Dumenci, Almqvist, Weintraub, Bilenberg et al., 2007; Rescorla, Achenbach, Ivanova, Dumenci, Almqvist, Bilenberg et al., 2007), several authors are convinced that as long as norm groups for specific populations with specific instruments are lacking, the assumption cannot be accepted that norms are omnicultural (Petot, Petot, & Achenbach, 2008). This is specifically relevant in the 21<sup>st</sup> century, characterised by new economic problems, new family structures, millions of immigrants, and blending of different subcultures within and between communities (Achenbach et al., 2008).

The most frequently used method in Flanders to screen for behavioural and emotional problems is Achenbach's Empirically Based Assessment system (ASEBA) (Schittekatte, Bos, Spruyt, Germeijs, & Stinissen, 2003). The system, which is based on observation of children in US epidemiological studies (Achenbach & Rescorla, 2001), comprises different versions of child, parent and teacher instruments, according to the age of the child, in order to allow for multi-informant assessment. For children between 6 and 18 years,

the Child Behaviour Checklist (CBCL; Achenbach & Rescorla, 2001) is one of the most important instruments of ASEBA and translated versions were adopted in more than 30 societies. Using parent reports on the preceding six months, it assesses 118 different emotional or behaviour problems. Items are scored on a three-point scale, allowing the calculation of broader-band scale scores. The scores are categorised in eight domain-specific syndrome scales as well as in two broad-band scales: Internalising and Externalising problems. The checklist provides normalised *T*-scores with *T*-scores  $\geq 64$  defined as high (clinical cutpoint) for Internalising problems, Externalising problems and for the Total Problems score, whereas for the eight domain-specific scales, *T*-scores  $\geq 70$  are defined as high (clinical cutpoint).

The validity and reliability of the Child Behaviour Checklist (CBCL, Achenbach, 1991) has been demonstrated across different cultures, including the Netherlands (De Groot, Koot, & Verhulst, 1994) and Belgium (Hellinckx, Grietens, & Verhulst, 1994). Concurrent validity is often demonstrated via a comparison with The Strengths and Difficulties Questionnaire (SDQ, Goodman, 1997). This is a rather new and brief measure, also aiming to screen for behavioural and emotional problems in children and adolescents. In a Flemish sample of 4-8 year olds, correlations between the SDQ and the CBCL were quite high and varied between .61 and .74 (Van Leeuwen *et al.*, 2006). A more recent Flemish study (Janssens & Deboutte, 2009), comparing parent versions of both instruments in a clinical sample, revealed correlations of .81 for the Total Problems score and the Externalising scale, and .70 for the Internalising scale, demonstrating excellent validity for the CBCL. While these concurrent validity findings are very promising, it must be noted that both studies employed the CBCL-1991 version. In the revised version of the CBCL (Achenbach & Rescorla, 2001) a (limited) number of items that were unscored or rare, were replaced with items that sharpen the assessment of important syndromes. Most countries recently switched over to the use of this (translated) CBCL-2001 version, but the psychometric properties of the Dutch CBCL-2001 version have never been studied in Flanders. Therefore, the first aim of the present study is to replicate previous positive findings concerning the reliability and concurrent validity of the CBCL-2001 version.

So far, only one study explored the construct validity of the 8-syndrome structure of the CBCL-1991 version in 30 different societies, including children from the Netherlands, France, Turkey, South-Korea as well as Flemish children (Hellinckx *et al.*, 1994; Ivanova *et al.*, 2007). Fit indices strongly supported the 8-syndrome structure in each of the 30 societies and provided first evidence that patterning of problem items assessed by the CBCL is similar in all societies tested. The authors concluded that although the CBCL was developed in the US and may assess problems that are particularly relevant

for US children, the data suggest a somewhat 'universal' syndrome structure tapping the diverse problems of all children. Furthermore, Rescorla et al. (2007) demonstrated additional support for the cross-cultural robustness of the CBCL. Based on the same dataset and comparing the means on all the different items, parents belonging to 31 different societies rated their children in a very similar way (with an overall 'normal' mean raw score of 22.5 on the Total Problems scale), with some societies score 1 *SD* higher or lower and 19 societies, including Belgium, within 1 *SD*. Although the robustness of the structure of the CBCL is quite convincing and provides a common language between cultures, we were more concerned whether the cutpoints of the CBCL 2001-version are not affected by small cultural differences and time trends. Therefore the second aim of the present study was to explore the usefulness of the existing US norms for both the broad-band scales as well as for the different syndrome scales for a representative sample of children and adolescents from the Flemish community.

To conclude, the objectives of the study are to examine some psychometric properties of the CBCL/1½-5 and CBCL/6-18 in a community sample of Flemish children between 1½ and 18 years of age. In addition, means, standard deviations and cutpoints will be analysed for the different age groups and for both sexes, which allows comparisons with findings in the US.

## Method

### *Participants*

A representative sample of Flemish children and their parents was invited to participate in this study (selection procedure: see below). Parents of 888 children, divided in 170 CBCL/1½-5 cases (84 boys and 86 girls) and 718 CBCL/6-18 cases (314 boys, 404 girls) participated. As Table 1 indicates, almost half of the total sample (46.40%) was between 6 and 11 years old. Families were excluded when the mother mastered only a foreign language, when the age of the child did not fall within the age range, when the child was diagnosed with a clinical DSM-disorder (e.g., dyslexia, ADHD, ...) or when families were in psychological treatment with their child.

Data collection took place via schools and day care centres between April 2009 and February 2010, following a standardised procedure and with respect to stratification according to four inclusion criteria: region, educational system, age and gender. The selection of schools and day care centres was guided by two important factors, namely region (province and level of urbanisation) and educational system of the child. About half of the schools agreed to participate in the study. In each school there was a random selection of five children per grade, with respect to a good distribution of different ages and

sexes<sup>[1]</sup>. The community sample was considered to be representative for children between 1½ and 18 years old of the Dutch speaking part of Belgium (i.e., the Flemish community).

### ***Instruments***

#### ***Achenbach system of empirically based assessment (CBCL/1½-5 and CBCL/6-18) (Achenbach & Rescorla, 2001)***

The Achenbach System of Empirically Based Assessment (ASEBA) is the most frequently used set of dimensional instruments to assess child psychopathology. In contrast to categorically based assessment, this dimensionally based assessment is characterised by a bottom-up approach, by obtaining scores for specific descriptors of children's functioning. Initial items were based on reviews of the clinical research literature and consultations with clinicians. Additionally, items were added or revised on the basis of findings in psychiatric case records and feedback from different kinds of informants and resulted in the final versions (Achenbach *et al.*, 2008). Two forms were of importance here.

*Child Behaviour Checklist for 1½ to 5 years: CBCL/1½-5.* The CBCL/1½-5 obtains information from parents or other guardians who know the child from a household context. The questionnaire asks about child functioning by means of 100 items that are scored on a 3-point scale ranging from 0 (not true) over 1 (somewhat or sometimes true) to 2 (very true or often true). The questionnaire was developed to assess the functioning of children during the last two months in a standardised format. Current norms of the CBCL/1½-5 are based on a national US sample of parents of children between the age of 18 (1 year and 6 months) and 71 months old (5 years and 11 months), without considerable physical or mental impairments and of whom the parents/guardians speak English. Children who experienced a stressful event or who were referred to mental health services or special education during the last 12 months were excluded from the group. The scales were finally normed on a sample of 700 children (362 boys and 338 girls). As gender and age differences were minimal in this age group, only one norm data set was obtained. After factor analysis on both norm groups, seven domain-specific syndrome scales were derived (*Emotionally reactive; Anxious/Depressed; Somatic complaints; Withdrawn; Sleep problems; Attention problems; Aggressive behaviour*), as well as two broad-band factors (*Internalising problems and Externalising problems*). Raw scores on the seven syn-

1. Comparison analyses, cutpoints and *T*-scores are available upon request (see also <http://www.testpracticum.ugent.be/>).

drome scales, on both broad-band factors as well as a Total Problems score can be transformed into normalised *T*-scores. Compared with the previous CBCL for this age group, the age range 2-3 years was elaborated, and spans now ages 1½-5 years.

*Child Behaviour Checklist for 6 to 18 years: CBCL/6-18.* The CBCL/6-18 obtains parents' reports of children's behavioural and emotional problems and competencies. Problems are assessed by means of 118 items, which are scored on a 3-point scale ranging from 0 (not true) over 1 (somewhat or sometimes true) to 2 (very true or often true), and 3 open-ended items. The current norms of the CBCL are based on a national US sample of 1753 children of children between 6 and 18 years, their parents and teachers. Different norms are available for girls and for boys, subdivided in two age groups (6-11 years and 12-18 years). Problem items are aggregated to provide scores on eight syndrome scales (*Anxious/Depressed; Withdrawn/Depressed; Somatic complaints; Social problems; Thought problems; Attention problems; Rule-breaking behaviour; Aggressive behaviour*) that have been derived through factor analytic methods. Additionally two intercorrelated broad-band factors (Internalising problems and Externalising problems) and a Total Problems score can be calculated. As is the case for the CBCL/1½-5 version, scores on the syndrome scales, broad-band factors and Total Problems score can be transformed into normalised *T*-scores. Finally, the revised CBCL/6-18 for school-aged children also contains DSM-IV oriented scales ([www.aseba.org](http://www.aseba.org)).

As Achenbach indicated in the 2001 Manual (Achenbach & Rescorla, 2001), there are many ways to look at construct validity. Regarding the internal structure of the Dutch CBCL, both confirmatory and exploratory factor analyses were already conducted in the Netherlands (De Groot et al., 1994; De Groot, Koot, & Verhulst, 1996). They showed identical consistencies compared with the exploratory and confirmatory factor analyses conducted on the original US version. Further evidence is provided by Achenbach and Rescorla (2007) with good fit indices for all 30 societies in the study, except Lebanon. These findings already strongly supported the cross-cultural robustness of the CBCL-syndrome structure.

#### *The strengths and difficulties questionnaires (SDQ)*

To study the concurrent validity of the ASEBA-instruments, Dutch versions of the SDQ – *Strengths and Difficulties Questionnaires* (SDQ, Goodman, 1997; Dutch translation: Treffers & van Widenfelt, 2000) (see van Widenfeldt, Goedhart, Treffers, & Goodman, 2003 for a description of the translation process), were used. Like the ASEBA, the SDQ assesses behavioural and emotional problems in youth and is available in different versions according

to age range and informant type (parents, teachers and youth). The SDQ queries 25 items divided into five scales (four negative scales and one positive scale: *Emotional symptoms; Conduct problems; Hyperactivity/Inattention; Peer relationship problems; Prosocial behaviour*). Each item is scored using a 3-point Likert-scale (0 = *not true*, 1 = *somewhat true* and 2 = *certainly true*). The scores for these subscales are compiled by adding the scores for the five corresponding items (some positively phrased items need to be recoded). The sum of the first four scales aggregates to provide a Total Problems score. The age range of the SDQ questionnaires is somewhat smaller (3-16 years) compared to the ASEBA questionnaires but the SDQ for children until 4 years differs only for 4 items compared with the 4-16 years version of the SDQ. This makes it easier to compare SDQ's between all age groups, compared with the ASEBA questionnaires CBCL/1<sup>1/2</sup>-5 and CBCL/6-18. Also, the SDQ is shorter than the CBCL and it includes items asking for positive qualities of the child. The psychometric properties of the Dutch version of the SDQ have been described in a Dutch (van Widenfeldt *et al.*, 2003) and Flemish (Van Leeuwen *et al.*, 2006) sample.

#### *Questionnaire recording demographic information*

A self-developed questionnaire recording demographic data of the child was distributed as well. This questionnaire probes for the following variables: degree of urbanisation/region, family situation, age of primary caregiver; current level of education of the child; mother tongue; highest level of education of the parents; professional activities of the parents; ethnicity of the parents; and some questions on possible contacts with treatment settings or mental health service use.

#### *Procedure*

The Ethical committee of Ghent University approved the protocol of this study. After giving consent to participate, the parents of the selected children were asked to complete the CBCL, as well as the SDQ and a self-developed questionnaire to record demographic information. Data collection took place via schools and day care centres between April 2009 and February 2010, following a standardised procedure. Most schools were visited by the project coordinator (Justine Callens) or by university students (both volunteering students or in exchange for course credits). The tests were administered as part of a larger project. Although multi-informant information was collected in the present study only findings on the CBCL and the SDQ are reported. The response rate was 33.5% for the CBCL/1<sup>1/2</sup>-5 and 67.1% for the CBCL/6-18. For 82.3% of the CBCL data, an SDQ equivalent was available. In 264 cases

(41.1%) questionnaires of both parents were available. In those cases, the questionnaire as completed by the mother was used in the analysis for this paper, to enhance homogeneity. Consequently, as in twenty-two cases (3.4%) only the fathers participated, we decided to exclude them here.

**Table 1**  
*The study group of Flemish children: Distribution of age and gender*

Age Group	Gender		Total
	Boys	Girls	
1½ - 05	84	86	170
06 - 11	183	229	412
12 - 18	131	175	306
<b>Total</b>	398	490	888

**Analyses**

Data analysis was performed using SPSS 16.0 and R 2.12.1. First, to obtain information on the internal consistency of the subscales, Cronbach's alpha values were calculated. Second, one-way ANOVA's were run to compare scale scores between boys and girls, and between different ages. Independent samples t-tests were performed to assess the deviance between the Flemish sample and the American norms.

It is generally acknowledged that raw scale scores can differ, depending on the version that is used (Achenbach & Rescorla, 2001, p. 166). Therefore, in clinical practice, the use of *T*-scores is preferred when interpreting the data, allowing also comparisons between scales with a different number of items and a different score distribution. For each age group (a) a clinical cutpoint for the domain specific syndrome scales is determined as the minimum raw score corresponding with *T*-score  $\geq 70$  or 98<sup>th</sup> percentile; while for the broad-band scales Internalising, Externalising and Total Problems score, this clinical cutpoint is determined as the minimum raw score corresponding to a *T*-score  $\geq 64$ , or 92<sup>nd</sup> percentile, (b) the borderline clinical range (e.g., range between clinical and subclinical cutpoint) is defined here as the raw scores corresponding with a *T*-score between 65 and 69 (93<sup>rd</sup> to 97<sup>th</sup> percentiles) for the domain specific syndrome scales and a *T*-score between 60 and 63 for the broad-band scales Internalising, Externalising and Total Problems score (84<sup>th</sup> to 90<sup>th</sup> percentiles) (Achenbach & Rescorla, 2001, p. 24-25).

Finally, the concurrent validity was evaluated by computing the Pearson product-moment correlation between equivalent CBCL and SDQ scales.



## Results

### *Scale reliability*

The internal consistency between the 100 items of the CBCL/1½-5 was estimated by calculating Cronbach's  $\alpha$  reliability coefficient. The internal consistency was high with  $\alpha = .94$ . The internal consistencies of the Internalising and Externalising scales are  $\alpha = .84$  and  $\alpha = .89$ , respectively. Also for the CBCL/6-18, the Cronbach's  $\alpha$  reliability coefficient was calculated. Similar to the CBCL/1½-5, the internal consistency between the 118 items was high with  $\alpha = .94$ . The internal consistencies of the Internalising and Externalising scales are both  $\alpha = .88$ .

### *Comparison of the Flemish data with US norms*

#### *Child behaviour checklist for 1½ to 5 years: CBCL/1½-5*

Table 2 presents an overview of the raw mean scores and standard deviations for the different subscales of the CBCL/1½-5 in both the Flemish sample ( $N = 170$ ) and the American sample ( $N = 700$ ) (Achenbach & Rescorla, 2001). Parallel to the American study of Achenbach and Rescorla (2007), the sample was not further divided in smaller subgroups based on gender and age category.

For 8 of the 10 scales, the raw mean scores of the Flemish sample differed significantly from those of American children. As can be seen in Table 2, Flemish children between one and a half and five years have *significantly lower* mean scores for all scales, except the syndrome scales Emotionally Reactive and Somatic complaints, for which no significant differences in mean scores were found.

In a second step, after calculating the percentiles, the equivalent cutpoints of the Flemish sample were compared to those of American children and added in Table 2. For 7 scales, the clinical cutpoint was lower, compared to American peers. Regarding the borderline clinical cutpoints, also 7 cutpoints were found to be lower. For example, for the broad-band scales Internalising and Externalising, the cutpoints differed 2 to 4 points (lower for Flemish children), while the borderline and clinical cutpoints for the Total Problems scale were even 12 and 13 points lower, respectively. Differences between cutpoints of the seven syndrome scales were less than 2 points, except for Aggressive behaviour, where there was a 3 point difference for the borderline cutpoint.

**Table 2**  
*Mean and standard deviations for raw scores and T-scores of the norm group CBCL/1<sup>1/2</sup>-5 years United States versus Flemish study; Cutpoints for each scale; t-tests for raw scores*

Scales		United States (N = 700)	Flemish <sup>a</sup> sample (N = 170)	Comparison U.S. and Flemish sample		
				t	df	Sig.
<i>Emotionally Reactive</i>	Raw score	2.4 (2.2)	2.4 (2.6)	0	231.19	1
	T-score	54.0 (5.7)	54.1 (6.2)			
	Sub clinical score	6	7+			
	Clinical score	9	11+			
<i>Anxious/Depressed</i>	Raw score	2.9 (2.3)	1.5 (1.7)	8.93	336.57	<.001
	T-score	54.2 (5.7)	53.9 (5.8)			
	Sub clinical score	7	5-			
	Clinical score	9	7-			
<i>Somatic Complaints</i>	Raw score	1.8 (1.9)	1.5 (1.7)	1.88	868	.060
	T-score	54.0 (5.8)	53.8 (5.8)			
	Sub clinical score	5	5			
	clinical score	7	6-			
<i>Withdrawn</i>	Raw score	1.5 (1.7)	0.8 (1.3)	5.90	324.97	<.001
	T-score	54.1 (5.8)	54.0 (6.0)			
	Sub clinical score	5	3-			
	Clinical score	6	5-			
<i>Sleep Problems</i>	Raw score	2.8 (2.4)	2.0 (2.3)	3.93	868	<.001
	T-score	54.2 (5.7)	54.1 (6.3)			
	Sub clinical score	8	6-			
	Clinical score	9	9			
<i>Attention Problems</i>	Raw score	2.5 (1.9)	1.9 (1.9)	3.70	868	<.001
	T-score	54.1 (5.6)	54.1 (6.6)			
	Sub clinical score	6	6			
	Clinical score	7	7			
<i>Aggressive Behaviour</i>	Raw score	10.4 (6.4)	8.5 (5.9)	3.52	867	<.001
	T-score	54.2 (6.0)	54.1 (6.5)			
	Sub clinical score	21	18-			
	Clinical score	25	23-			
<i>Internalising</i>	Raw score	8.6 (6.2)	6.3 (5.9)	4.38	868	<.001
	T-score	50.0 (9.9)	50.1 (9.7)			
	Sub clinical score	14	12-			
	Clinical score	18	14-			
<i>Externalising</i>	Raw score	12.9 (7.7)	10.4 (7.1)	3.84	867	<.001
	T-score	50.0 (9.9)	50.0 (9.9)			
	Sub clinical score	21	17-			
	Clinical score	25	22-			
<i>Total Problems</i>	Raw score	33.3 (18.7)	24.7 (17.3)	5.44	867	<.001
	T-score	50.1 (9.9)	50.0 (9.9)			
	Sub clinical score	52	40-			
	Clinical score	61	48-			

a Note column 4: - and + refer to the difference between cut-offs (United States vs. Flemish sample)

*Gender and age differences for the CBCL/1<sup>1/2</sup>-5*

Significant differences were found between boys and girls for four of the seven syndrome scales (Emotionally Reactive:  $F(1, 168) = 4.07, p = .045$ ; Withdrawn:  $F(1, 168) = 5.17, p = .024$ ; Attention Problems:  $F(1, 168) = 11.47, p = .001$ ; Aggressive Behaviour:  $F(1, 167) = 8.55, p = .004$ ) and for the broad-band scales (Externalising:  $F(1, 167) = 11.24, p = .001$ ; Internalising:  $F(1, 168) = 4.15, p = .043$ ) and Total Problems score:  $F(1, 167) = 6.35, p = .013$ . In all cases the mean scores for boys were higher. No significant age differences were found, except for the subscale Attention Problems,  $F(4, 165) = 3.01, p = .020$ , with the highest mean on this scale observed for the 1-year-olds.

*Child behaviour checklist for 6 to 18 years: CBCL/6-18*

Tables 3 and 4 present an overview of the raw mean scores and standard deviations for the different subscales of the CBCL/6-18 in both the Flemish sample ( $N = 718$ ) and the American sample ( $N = 1753$ ) (Achenbach & Rescorla, 2001) for boys and girls respectively. Parallel to the American study of Achenbach and Rescorla (2007), the sample was divided in four different groups, based on gender and age category (6-11 and 12-18 years).

Comparing the raw mean scores of the Flemish and American boys between 6 and 11 years on the CBCL/6-18, 6 of the 11 scales were significantly different (Table 3), whereby for 3 scales Flemish boys scored higher (including the Internalising scale) and for 3 other scales, their mean scores were lower (including the Externalising scale). In a second step, after calculating the percentiles, the cutpoints were determined (11 clinical cutpoints and 11 borderline clinical cutpoints) of the Flemish sample and they were compared to those of American children (not included in the table<sup>[2]</sup>). Visual inspection revealed no consistent pattern. For example, the cutpoints for the Externalising scale reveal that Flemish boys scored 2 points lower on both the borderline and clinical cutpoints, while for the Internalising scale, Flemish boys scored 3 (borderline clinical cutpoint) to 5 points (clinical cutpoint) higher.

For boys between 12 and 18 years, 3 of the 11 scale means were significantly different between Flemish boys and their American peers (see Table 3). Mean scores were significantly lower for the Flemish sample on the subscales Rule-breaking Behaviour, Aggressive Behaviour, and Externalising problems. After calculating the percentiles, the cutpoints were determined.

2. The cutpoints for the 6-18-year-olds, based on our research, can be consulted via <http://www.testpracticum.ugent.be/>

**Table 3**  
*Mean en standard deviations for the raw scores of the norm group CBCL/6-18 years United States versus Flemish sample – boys; t-tests for raw scores*

Samples	Gender Age	United States		Flanders		Comparison U.S. and Flemish samples						
		6-11 N = 387	12-18 N = 527	6-11 N = 183	12-18 N = 131	t	df	Sig.	t	df	Sig.	
<i>Scales</i>												
<i>Anxious/Depressed</i>	Mean and (SD) raw score	2.8 (2.7)	2.6 (2.7)	3.4 (3.9)	3.0 (3.6)	-1.87	265.47	.062	-1.19	166.50	.237	
<i>Withdrawn/Depressed</i>	Mean and (SD) raw score	1.1 (1.6)	1.9 (2.2)	1.4 (1.8)	1.5 (2.0)	-2.01	568	.045	1.89	655	.059	
<i>Somatic Complaints</i>	Mean and (SD) raw score	1.1 (1.7)	1.1 (1.8)	1.5 (2.0)	1.3 (1.9)	-2.33	305.91	.021	-1.12	653	.265	
<i>Social Problems</i>	Mean and (SD) raw score	2.4 (2.6)	1.8 (2.3)	2.0 (2.5)	1.7 (2.4)	1.74	568	.083	0.44	656	.659	
<i>Thought Problems</i>	Mean and (SD) raw score	1.8 (2.0)	1.6 (1.9)	1.8 (2.1)	1.6 (1.9)	0.00	567	1.0	0.00	656	1.0	
<i>Attention Problems</i>	Mean and (SD) raw score	3.8 (3.4)	4.0 (3.7)	3.4 (3.3)	3.6 (3.4)	1.32	568	.186	1.12	656	.261	
<i>Rule-Breaking Behaviour</i>	Mean and (SD) raw score	1.9 (2.1)	2.8 (3.4)	1.2 (1.5)	1.8 (2.1)	4.55	480.92	<.001	4.24	320.93	<.001	
<i>Aggressive Behaviour</i>	Mean and (SD) raw score	4.7 (4.3)	4.7 (4.8)	3.9 (4.4)	3.5 (3.9)	2.06	568	.040	3.00	237.99	.003	
<i>Internalising</i>	Mean and (SD) raw score	5.1 (4.8)	5.6 (5.3)	6.3 (6.5)	5.8 (6.2)	-2.23	279.21	.027	-0.34	178.28	.735	
<i>Externalising</i>	Mean and (SD) raw score	6.6 (6.0)	7.5 (7.5)	5.1 (5.5)	5.3 (5.6)	2.85	559	.005	3.74	259.04	<.001	
<i>Total Problems</i>	Mean and (SD) raw score	23.4 (16.9)	23.7 (19.0)	21.3 (18.0)	20.4 (17.5)	1.36	568	.176	1.81	656	.071	

**Table 4**  
*Mean en standard deviations for the raw scores of the norm group CBCL/6-18 years United States versus Flemish sample – girls; t-tests for raw scores*

Samples	Gender Age	United States		Flanders		Comparison U.S. and Flemish sample				
		6-11 N = 390	12-18 N = 449	6-11 N = 229	12-18 N = 175	6-11 df	t	Sig.	12-18 df	Sig.
<b>Scales</b>										
<i>Anxious/Depressed</i>	Mean en (SD) raw score	3.2 (2.9)	3.2 (3.1)	3.0 (3.3)	3.0 (3.4)	429.85	0.76	.447	622	.482
<i>Withdrawn/Depressed</i>	Mean en (SD) raw score	1.4 (1.7)	1.9 (2.1)	1.1 (1.8)	1.7 (2.1)	617	2.07	.039	622	.286
<i>Somatic Complaints</i>	Mean en (SD) raw score	1.3 (1.7)	1.4 (1.9)	1.3 (1.9)	1.5 (2.0)	616	0.00	1.0	619	.563
<i>Social Problems</i>	Mean en (SD) raw score	2.6 (2.6)	1.8 (2.3)	2.1 (2.6)	1.5 (1.9)	617	2.31	.021	381.22	.096
<i>Thought Problems</i>	Mean en (SD) raw score	1.7 (1.8)	1.4 (1.7)	1.6 (2.0)	1.7 (2.2)	617	0.64	.522	258.97	.105
<i>Attention Problems</i>	Mean en (SD) raw score	3.2 (3.1)	2.7 (3.1)	3.2 (3.1)	3.2 (3.2)	617	0.00	1.0	622	.073
<i>Rule-Breaking Behaviour</i>	Mean en (SD) raw score	1.6 (1.8)	2.2 (3.0)	1.3 (1.6)	1.7 (2.0)	523.38	2.15	.032	472.01	.016
<i>Aggressive Behaviour</i>	Mean en (SD) raw score	4.5 (4.3)	4.4 (4.7)	4.0 (4.6)	3.5 (3.8)	617	1.36	.174	389.56	.014
<i>Internalising</i>	Mean en (SD) raw score	6.0 (5.0)	6.5 (5.7)	5.4 (6.0)	6.1 (6.2)	411.73	1.28	.203	622	.443
<i>Externalising</i>	Mean en (SD) raw score	6.1 (5.6)	6.6 (7.0)	5.3 (5.8)	5.1 (5.4)	617	1.69	.091	408.50	.005
<i>Total Problems</i>	Mean en (SD) raw score	22.9 (16.6)	22.0 (18.2)	20.1 (18.0)	20.1 (18.0)	617	1.96	.050	622	.240

Inspection of the cutpoints for all scales revealed no consistent pattern: 5 cutpoints were identical, 6 were higher and 11 were lower for the Flemish boys (not included in the table). For example, while the borderline and clinical cutpoint scores for the scale Internalising problems were only one point higher for the Flemish boys, the Externalising problems and the Total Problems scale revealed cutpoints which were substantially lower (for the borderline clinical cutpoints, the difference was 4 to 7 points and for the clinical cutpoints, the differences were 5 and 2 points, respectively) compared to American children.

Scale comparisons indicated further that also *Flemish girls between 6 and 11 years* scored *significantly lower* than their American peers for 4 of the 11 scales (see Table 4): the scales Withdrawn/Depressed, Social Problems, Rule-breaking Behaviour and also the Total Problems scale. After calculating the percentiles, the cutpoints were determined. Visual inspection revealed that 10 cutpoints for the different scales for Flemish girls between 6 and 11 years were lower (not included in the table), while only 4 were higher. No differences were found for the cutpoints on the Internalising scale while the cutpoints for the Externalising scale were 2 points lower for the Flemish girls. Differences in the cutpoints for the Total Problems scale were substantial, with Flemish girls having (both clinical and borderline) cutpoints that are 4 points lower.

Finally, the means of the Flemish sub sample of *girls between 12 and 18 years* differed on 3 of the 11 subscales from the American norm group (see Table 4). For all three scales, including the Externalising scale, Flemish girls scored *significantly lower*. Further (not included in the table) after calculating the percentiles and cutpoints, the cutpoints for 3 scales (Anxious/Depressed, Thought Problems and Attention Problems) were higher, while for 4 scales (Social Problems, Rule-breaking Behaviour, Aggressive Behaviour and Externalising) the cutpoint scores were lower for Flemish girls (12-18 years). While there were no differences in cutpoint scores for the Internalising scale and Total Problems scales, the difference for the Externalising scale was 2 points.

#### *Gender and age differences for the CBCL/6-18*

No significant differences between boys and girls could be found in the mean subscale scores. When comparing the subscale means between age categories, a significant difference could be found for 3 of the 11 subscales: for Social Problems, children between 6 and 11 years score higher compared to the group between 12 and 18 ( $F(1, 716) = 6.42, p = .012$ ), while for the subscales Withdrawn/Depressed ( $F(1, 716) = 7.70, p = .006$ ) and Rule-breaking Behaviour ( $F(1, 716) = 12.66, p < .001$ ), the older age group showed significantly more problems.

**Concurrent validity: comparison of the CBCL with the SDQ**

The Pearson product-moment correlations between the equivalent scales of the CBCL and the SDQ were calculated (*Total Problems; Internalising/Emotional Symptoms; Externalising/Conduct Problems; Attention Problems/Hyperactivity/Inattention; Social Problems (CBCL/ 6-18)/Peer relationship problems*) (see Table 5). All correlations were significant at  $p < 0.01$  level. In general, correlations with the SDQ seemed to be stronger for the CBCL/1½-5 compared to the CBCL/6-18. The correlations between 1½-5 sample, the CBCL/6-18 sample and SDQ scales were substantial for all scales except for *Social Problems/Peer relationship problems* ( $r = .60$ ).

**Table 5**  
*Correlations between equivalent CBCL and SDQ scales*

Equivalent scales (CBCL/SDQ)	Sample CBCL/1½-5, respondent =mother (N = 170)		Sample CBCL/6-18, respondent = mother (N =718)	
	N	r	N	r
<i>Total Problems score / Total SDQ</i>	82	.80**	554	.80**
<i>Internalising / Emotional Symptoms</i>	82	.75**	554	.75**
<i>Externalising / Conduct Problems<sup>a</sup></i>	25	.90**	554	.68**
<i>Attention Problems / Hyperactivity/Inattention</i>	82	.77**	554	.74**
<i>Social Problems/ Peer relationship problems<sup>b</sup></i>	-	-	554	.60**

\*\*  $p < 0.01$

- a Only the 5 year olds were considered, because of a difference in version of the SDQ for the youngest and older children
- b Scale does not exist in the SDQ for the youngest age group

**Discussion**

ASEBA is a widely used battery of instruments indicative for screening of psychopathology in children and adolescents. Specifically the CBCL is one of the most popular measures, also in the Flemish community. However, the psychometric properties of the CBCL were based on studies in US samples and until now, the scientific usefulness of the most recently translated version was never questioned. In this study the reliability and concurrent validity of the Dutch CBCL-2001 as well as the usefulness of the US norms were evaluated in two different community samples (1½-5 and 6-18 years) of children and adolescents.

A first criterion in evaluating the usefulness of a screening instrument is the quality of its internal consistency. Results of the study provided strong support for the reliability of both Dutch versions of the CBCL. The findings are comparable with alpha coefficients found in international studies, ranging between .90 and .96 (Rescorla et al., 2007) and with those in a Flemish study in a clinical sample with a previous CBCL version (Janssens & Deboutte, 2009).

Evidence was also found for the construct validity of the Dutch CBCL. Comparing the CBCL subscales with the conceptually related SDQ subscales revealed that all correlations were significant. For the Total Problems score a robust correlation was found of .80 for both the sample 1½-5 years and the sample 6-18 years whereas for the Internalising and Externalising scale the correlations varied between .68 and .90. Correlations between the subscales of the CBCL with the SDQ seem to be even stronger for the CBCL 1½-5 than for the CBCL/6-18. Compared to another Flemish study, in 4-7 years olds (Van Leeuwen et al., 2006), the present study revealed even better SDQ-CBCL correlations for the youngest age group. It was assumed that specifically for older children, parent reports can differ according to the specific items that were questioned. Also comparable to our study, significant correlations between the (1991 version of the) CBCL and the SDQ were found in a Flemish study in a clinical sample with a broad age range (3-18 years), with correlations varying between .70 (Internalising scale) and .81 (Externalising and Total scale) (Janssens & Deboutte, 2009). Probably, in clinical samples, heterogeneity can explain the somewhat lower correlations.

When comparing the mean raw scores of the different subsamples with the US norms some important differences emerged. Generally spoken, cut-points for the Total Problems scale and the Externalising scale and subscales, were lower in the Flemish sample. The majority of these differences were supported by a significant difference in the mean scores between norm groups. On the CBCL/6-18, Flemish children in this study had a raw mean score on the Total Problems scale of respectively 20.1 (girls 6-11 years), 20.1 (girls 12-18 years), 21.3 (boys 6-11 years) and 20.4 (boys 12-18 years) which differ 2-3 points with the US norms. While Crijnen et al. (1997) already indicated that Belgian children scored below the omnicultural mean of 22.4 (CBCL-1991 version), Rescorla et al. (2007) stated that such difference should not be seen as problematic, as long as they are less than 1 *SD*. However, visual inspection of the differences in cutpoints on the different scales of both the CBCL/6-18 version and the CBCL/1½-5 version in the current study, revealed that according to US borderline clinical and clinical cutpoints, a substantial part of the Flemish children will not be identified as 'at risk children', suggesting an underestimation. Consequently, this can lead to more false negative cases, compared with what is generally expected when using the CBCL.



It must be acknowledged that for the different scales, somewhat more differences with US norms were found for Flemish boys compared with Flemish girls. Interestingly here, for the Internalising scale of the CBCL/6-18 the norm and cutpoints for boys 6 to 11 years were not lower but substantially higher. So, specifically for this subsample, Flemish parents report significantly more internalising problems in their own children, compared to American parents.

Some important age differences were noticed, which justifies the suggested use of minimally 3 age norm groups (1½-5 years, 6-11 years and 12-18 years). Although in this study for the CBCL/6-18 years, no significant differences could be found in the mean subscale scores for boys and girls, the different cutpoints in the subsamples plaid for the relevance of different norms for boys and girls. Remarkably, and in contrast with our findings, Achenbach and Rescorla (2001) did not find evidence for splitting the 1½-5 years sample in gender-specific groups. Until now, only a few studies explored gender differences in the 1½-5 years sample and the findings seem rather inconsistent. However, a number of studies have shown higher rates of externalising problems in boys (e.g., Gross, Fogg, Young, Ridge, Cowell, Richardson *et al.*, 2006). The authors therefore strongly recommend further exploring the necessity of gender specific norms in all age groups in future research.

The present study has some important clinical implications. In recent years, western countries have experienced several societal changes. Research further revealed that mental health problems in children slightly increased over recent decades (Tick *et al.*, 2007). Therefore, one must be aware that identifying a particular child as 'clinical' depends on the norm group it is compared with. Given the sparse number of studies that compared identical assessments of population samples at different time points, the study of Tick *et al.* (2007) is unique because of an interesting comparison between CBCL-parent reports collected in 1983 and 20 years later. In their conclusion the authors reported that only small meaningful differences were found. Changes were strongest between 1993 and 2003 and mainly concerned internalising problems. Although explaining time trends is difficult, these results indicate that using 1991 norms or 2001 norms will affect our clinical work.

Besides the time factor, also other factors may require the availability of accurate, cultural specific norms. While each individual parent holds different thresholds and personal standards when rating problem behaviour (van der Ende, 1999), a fairly consistent finding in literature is that parents of minority groups score higher than parents of white Caucasian children on parent-reported behaviour problem measures, which raises questions about the robustness of an omnicultural norm group (Gross *et al.*, 2006). Although the abovementioned studies can be criticised for the fact that they confuse ethnic

background with income and socio-economic status, the findings of the current study indicate that it is worthwhile to consider specific subcultures.

Several explanations can be found for the differences found between cut-point scores in the Flemish sample compared to the US sample. First, within the Flemish community, there is a high prevalence of middle and high-income groups (Reynders, Nicaise, & Van Damme, 2005), which may account for a lower rate of child problems reported by parents compared to more heterogeneous cultures like in the US. Second, several studies have illustrated that different aspects of parenting (including beliefs, goals, values and behaviour style) are at least partly culturally constructed (Boushel, 2000; Rubin & Chung, 2006). This is especially relevant for small differences among societies within the same cultural area, such as the Western world. While it is generally believed that there are no real cultural differences among societies in Europe, the US and other parts of the European Diaspora, the findings in the present study may refer to the fact that parents in different cultures *think* in a slightly different way about their children, and consequently also report differently about child problems. Harkness and Super (2006) speak about 'parental ethnotheories' in this context, to explain the common themes as well as the variations they found in the discourses of native-born, middle-class parents from six different cultural samples in Western countries.

Although the CBCL is a well-validated instrument for screening purposes, one must be careful in adopting it. The CBCL can by no means be used as a diagnostic tool (Warnick, Weersing, Scahill, & Woolston, 2009). The extent to which the CBCL succeeds in an accurate identification of 'true' cases in relation to clinical diagnoses has been described in terms of sensitivity and specificity. To make it easy for clinicians and researchers to judge children's functioning, each scale score of a screening instrument is only relevant when it refers to percentiles and *T*-scores that enable to compare a specific child's raw score with normative samples. For the specific scale scores of the CBCL, a borderline clinical range refers to the 93<sup>rd</sup> till 97<sup>th</sup> percentile; lower scores refer to the normal range, while scores above the 97<sup>th</sup> percentile were defined as clinical and indicate that the person who filled in the questionnaire reported enough problems of clinical concern. So, cutpoints were important markers for identifying an at risk child. Also, although Achenbach suggests that a *T*-score on the broad-band scales greater than 63 is generally indicative of problem behaviour, specificity and sensitivity still vary across different studies, with a generally estimated sensitivity of .66 and an estimated specificity of .83, indicating that numbers of false positives and false negatives are still quite large. The use of a 'multiple stage' strategy in the assessment of psychopathology has been recommended (Kendall, Cantwell, & Kazdin, 1989). This approach involves the use of a screening test, such as a questionnaire, to select potential cases for further assessment by means of a cutpoint score (which

meets adequate criteria of sensitivity and specificity). A next assessment period for those participants exceeding the cutpoint score involves a second administration of the questionnaire along with a structured interview to assess Diagnostic and Statistical Manual for Mental Disorders (DSM-IV; APA, 1994) diagnoses. By following this strategy, not only the core symptoms of psychopathology are assessed, but also the duration of a particular constellation of symptoms, as well as the severity and interference with daily functioning.

Careful attention was given to compose a study sample that met most criteria for representativeness. Besides an equal distribution between both sexes and a sufficient number of participants in all age groups, comparison analyses (see footnote 1) with recent norm data for the Flemish community also reveal that also both regional spreading and level of urbanisation largely correspond to the proportions in the Flemish Community, except for the Eastern region (Limburg) (Source: Studiedienst Vlaamse Regering; FOD: Economy – Department of Statistics, Population statistics: Population of children between 0 and 17 years in Flanders). With respect to the social economical status of the families based on job status of the mother (Source: Reynders *et al.*, 2005, LOA-report n° 31) and relying on the schema classification of Erickson, Goldthorpe, and Portocarero (1979) the study sample can be considered as representative especially for the lower social class (20% of the sample). However, compared to statistics collected in 2002, the upper SES class participated slightly less in the 1½-5 years sample and slightly more in the 6-18 years sample. This skewed distribution can however partially be explained by a non-classified rest group who preferred to remain anonymous. Finally, also representativeness for family composition was explored, based on comparisons with PSBH (Panel Studie van Belgische Huishoudens) 2002 statistics (Source: Koning Boudewijnstichting, 2008). The findings reflect a good distribution of the different families according to their current composition and are comparable to what can be expected now in the Flemish community. To conclude, the random sample met most criteria for representativeness although some caution is warranted regarding SES.

Several limitations to this study can be noted. First, the described psychometric properties were based on only one Flemish community sample of children and replication or extending the study by doubling the sample size is warranted. Although much effort was made to compose a random sample and although the sample met different criteria regarding representativeness, the authors failed to show that SES distribution is completely conform Flemish community statistics as SES data were somewhat blurred by strict privacy demands. Also, specifically recruiting families for the youngest children of the CBCL/1½-5 years outside schools was difficult. Also other scholars reported low response rates for this age group (Rescorla *et al.*, 2007). Never-

theless, also in this sample, representativeness according to our criteria was more or less satisfying.

The low correlations found in other studies between different mother-father versions further indicated that child problems are not effectively captured by only one informant (Achenbach, McConaughy, & Howell, 1987; van der Ende, 1999) and all informants can provide a unique contribution to the assessment (Grietens, Onghena, Prinzie, Gadeyne, Van Assche, Ghesquière et al., 2004). In order to enhance the homogeneity of our norm group, we preferred to analyse the mother reports for all participants. It is not totally clear how many mothers and fathers were involved in the US norms and whether parental US norms were different when analysed separately for father or mother reports. In future research, the authors recommend to explore potential differences between two informants (father versus mother, parent versus teacher and parent versus child) and the surplus value of consulting multiple informants in screening psychopathology. Finally, the utility of the CBCL in clinical children remains also to be investigated.

To conclude, it must be recognised that careful screening of mental health problems in children affects the responsibility of each clinician and that it is only possible when the screening instrument is carefully chosen, based on its psychometric properties. Despite its limitations, the present study indicates that the translated Dutch version of the CBCL has promising characteristics although it can be questioned whether at the moment good and reliable norms are available.

### References

- Achenbach, T.M. (1991). *Manual for the child behavior checklist/4-18 and 1991 profile*. Burlington, VT: University of Vermont, Department of Psychiatry.
- Achenbach, T.M., Becker, A., Döpfner, M., Heiervang, E., Roessner, V., Steinhausen, H.C. et al. (2008). Multicultural assessment of child and adolescent psychopathology with ASEBA and SDQ instruments: Research findings, applications, and future directions. *Journal of Child Psychology and Psychiatry*, 49, 251-275.
- Achenbach, T.M., McConaughy, S.H., & Howell, C.T. (1987). Child adolescent behavioral and emotional problems – Implications of cross-informant correlations for situational specificity. *Psychological Bulletin*, 101, 213-232.
- 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.
- Achenbach, T.M., & Rescorla, L.A. (2007). *Multicultural supplement to the manual for the ASEBA school-age forms & profiles*. Burlington, VT: University of Vermont, Research Center for Children, Youth & Families.
- American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders (4th Edition)*. Washington, DC: American Psychiatric Association (APA).

- Boushel, M. (2000). Childrearing across cultures. In M. Boushel, M. Fawcett, & J. Selwyn (Eds.), *Focus on early childhood: Principles and realities* (pp. 65-77). Oxford: Blackwell Science.
- Braet, C., & van Aken, M.A.G. (2006). Developmental psychopathology: Substantive, methodological and policy issues. *International Journal of Behavioral Development, 30*, 2-4.
- Burns, B., Philips, S., Wagner, R., Barth, R., Kolko, D., Campbell, Y., & Landsverk, J. (2004). Mental health need and access to mental health services by youths involved with child welfare: A national survey. *American Academy of Child and Adolescent Psychiatry, 43*, 960-970.
- Costello, E.J., & Angold, A.C. (2000). Developmental epidemiology. In A.J. Samerpoint, M. Lewis, & L. Miller (Eds.), *Handbook of developmental psychopathology*, Second Edition. New-York: Plenum Publishers.
- Crijnen, A.A.M., Achenbach, T.M., & Verhulst, F.C. (1997). Comparisons of problems reported by parents of children in 12 cultures: Total problems, externalizing and internalizing. *Journal of the American Academy of Child and Adolescent Psychiatry, 36*, 1269-1277.
- Culbertson, J.L. (1993). Clinical child psychology in the 1990s: Broadening our scope. *Journal of Clinical Psychology, 22*, 116-122.
- De Groot, A., Koot, H.M., & Verhulst, F. (1994). Cross-cultural generalizability of the child behavior checklist cross-informant syndromes. *Psychological Assessment, 6*, 225-230.
- De Groot, A., Koot, H.M., & Verhulst, F. (1996). Cross-cultural generalizability of the youth self report and teacher's report form cross-informant syndromes. *Journal of Abnormal Child Psychology, 24*, 651-664.
- Erikson, R.J., Goldthorpe, J.H., & Portocarero, L. (1979). Intergenerational class mobility in three Western European societies. *British Journal of Sociology, 30*, 415-441.
- Goodman, R. (1997). The strengths and difficulties questionnaire: A research note. *Journal of Child Psychology and Psychiatry, 38*, 581-586.
- Grietens, H., Onghena, P., Prinzie, P., Gadeyne, E., Van Assche, V., Ghesquière, P., & Hellinckx, W. (2004). Comparison of mothers', fathers', and teachers reports on problem behavior in 5- to 6-year-old children. *Journal of Psychopathology and Behavioral Assessment, 26*, 137-146.
- Gross, D., Fogg, L., Young, M., Ridge, A., Cowell, J.M., Richardson, R., & Sivan, A. (2006). The equivalence of the child behavior checklist across parent race/ethnicity, income level, and language. *Psychological assessment, 18*, 313-323.
- Harkness, S., & Super, C.M. (2006). Themes and variations: Parental ethnotheories in Western cultures. In K.H. Rubin & O.B. Chung (Eds.), *Parenting believes, behaviors and parent child relationships* (pp. 61-80). New York: Taylor & Francis Group.
- Hellinckx, W., Grietens, H., & Verhulst, F. (1994). Competence and behavioral problems in 6-year-old to 12-year-old children in Flanders (Belgium) and Holland – A cross-national comparison. *Journal of Emotional and Behavioral Disorders, 2*, 130-142.

- Ivanova, M.Y., Achenbach, T.M., Dumenci, L., Almqvist, F., Weintraub, S., Bilenberg, N. et al. (2007). Testing the 8-syndrome structure of the child behavior checklist in 30 societies. *Journal of Clinical Child and Adolescent Psychology*, 36, 405-417.
- Janssens, A., & Deboutte, D. (2009). Screening for psychopathology in child welfare: The strengths and difficulties questionnaire (SDQ) compared with the Achenbach system of empirically based assessment (ASEBA). *European Child and Adolescent Psychiatry*, 18, 691-700.
- Jensen, A.L., & Weisz, J.R. (2002). Assessing match and mismatch between practitioner-generated and standardized interview-generated diagnoses for clinic-referred children and adolescents. *Journal of Consulting and Clinical Psychology*, 70, 158-168.
- 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.
- Koning Boudewijnstichting (2008). Samen werken aan een betere samenleving. *Het kind in een nieuw samengesteld gezin*, November 2008, 74.
- Oosterlaan, J., Baeyens, D., Scheres, A., Antrop, I., Roeyers, H., & Sergeant, J.A. (2008). *Handleiding bij de vragenlijst voor gedragsproblemen bij kinderen van 6 tot en met 16 jaar (VvGK6-16) [Manual for the questionnaire for behavioural disorders in children from 6 to 16 years]*. Amsterdam, The Netherlands: Harcourt Test Publishers.
- Petot, D., Petot, J.M., & Achenbach, T.M. (2008). Behavioral and emotional problems of Algerian children and adolescents as reported by parents. *European Journal of Child and Adolescent Psychiatry*, 17, 200-2008.
- Rescorla, L.A., Achenbach, T.M., Ivanova, M.Y., Dumenci, L., Almqvist, F., Bilenberg, N. et al. (2007). Behavioral and emotional problems reported by parents of children ages 6 to 16 in 31 societies. *Journal of Emotional and Behavioral Disorders*, 15, 130-142.
- Reynders, T., Nicaise, I., & Van Damme, J. (2005). *Longitudinaal onderzoek in het basisonderwijs: De constructie van een SES-variabele voor het SIBO-onderzoek. LOA-rapport*, 31. Leuven: Steunpunt loopbanen doorheen onderwijs naar arbeidsmarkt.
- Rubin, K.H., & Chung, O.B. (2006). *Parenting believes, behaviors and parent child relationships*. New York: Taylor & Francis Group.
- Schittekatte, M., Bos, A., Spruyt, K., Germeijs, V., & Stinissen, H. (2003). Rondvraag naar het diagnostisch instrumentarium en de noden in Vlaanderen [Reports and queries about the diagnostic instrumentarium and the needs in Flanders]. *Tijdschrift voor Orthopedagogiek, Kinderpsychiatrie en Klinische Kinderpsychologie*, 28, 50-62.
- Studiedienst Vlaamse Regering (2008). Populatie kinderen tussen 0 en 17 jaar in Vlaanderen [Population children between 0 and 17 years old in Flanders]. FOD: Economie – Afdeling Statistiek, Bevolkingsstatistiek. Retrieved January, 7th, 2010, from <http://www4.vlaanderen.be/dar/svr/Cijfers/Pages/Excel.aspx>

- Tick, N.T., van der Ende, J., & Verhulst, F.C. (2007). Twenty-year trends in emotional and behavioral problems in Dutch children in a changing society. *Acta Psychiatrica Scandinavica*, *116*, 473-482.
- Timbremont, B., Braet, C., & Roelofs, J. (2008). *Handleiding children's depression inventory (herziene versie) [Manual children's depression inventory (revised version)]*. Amsterdam, The Netherlands: Pearson Assessment and Information, B.V.
- van der Ende, J. (1999). Multiple informants: multiple views. In H.M. Koot, A.A.M. Crijnen, & R.F. Ferdinand (Eds.), *Child psychiatric epidemiology: Accomplishments and future directions* (pp. 39-52). Assen: Van Gorcum.
- Van Leeuwen, K., Meerschaert, T., Bosmans, G., De Medts, L., & Braet, C. (2006). The strengths and difficulties questionnaire in a community sample of young children in Flanders. *European Journal of Psychological Assessment*, *22*, 189-197.
- van Widenfeldt, B.M., Goedhart, A.W., Treffers, P.D.A., & Goodman, R. (2003). Dutch version of the strengths and difficulties questionnaire (SDQ). *European Child and Adolescent Psychiatry*, *12*, 281-289.
- Warnick, E.M., Weersing, V.R., Scahill, L., & Woolston, J.L. (2009). Selecting measures for use in child mental health services: A scorecard approach. *Administration and Policy in Mental Health and Mental Health Services Research*, *36*, 112-122.

Received February 22, 2010  
Revision received January 14, 2011  
Accepted June 28, 2011