



Published in final edited form as:

*J Abnorm Child Psychol.* 2009 August ; 37(6): 777–791. doi:10.1007/s10802-009-9320-6.

## Are there stable factors in girls' externalizing behaviors in middle childhood?

Rolf Loeber, Dustin Pardini, Magda Stouthamer-Loeber, Alison E. Hipwell, and Mark Semboer

### Abstract

Relatively little is known about the factor structure of disruptive behavior among pre-adolescent girls. The present study reports on exploratory and confirmatory factor analyses of disruptive girl behavior over four successive data waves as rated by parents and teachers in a large, representative community sample of girls ( $N = 2,451$ ). Five factors were identified from parent ratings (oppositional behavior/conduct problems, inattention, hyperactivity/impulsivity, relational aggression, and callous-unemotional behaviors), and four factors were identified derived from teacher ratings (oppositional behavior/conduct problems/callous-unemotional behaviors, inattention, hyperactivity-impulsivity, and relational aggression). There was a high degree of consistency of items loading on equivalent factors across parent and teacher ratings. Year-to-year stability of factors between ages 5 and 12 was high for parent ratings ( $ICC = .70$  to  $.88$ ), and slightly lower for teacher ratings ( $ICC = .56$  to  $.83$ ). These findings are discussed in terms of possible adjustment to the criteria for children's disruptive behavior disorders found in the Diagnostic and Statistical Manual for Mental Disorders.

---

Studies on disruptive behaviors in girls lag behind those on boys, and the question of what are appropriate symptom clusters of disruptive behavior in girls at a young age remains largely unanswered (e.g., Frick et al., 1993; Vaillancourt, Brendgen, Boivin & Tremblay, 2003). Most studies have assumed that symptom clusters of disruptive behavior are similar for each gender but that symptom frequency and severity tends to be lower for girls than boys (e.g., Loeber & Schmalting, 1985). These studies largely focused on symptoms associated with oppositional defiant disorder (ODD), conduct disorder (CD), and attention-deficit/hyperactivity disorder (ADHD) (e.g., Frick et al., 1993; Storvoll, Wichstrøm, Kolstad, & Pape, 2002). However, in recent years supplementary approaches have stressed indirect forms of aggression and callous-unemotional behaviors (e.g., Prinstein, Boegers, & Vernberg, 2001; Vaillancourt et al., 2003; Vaillancourt, Miller, Fagbemi, Côté, & Tremblay, 2007), which are symptoms that typically are not part of routine psychiatric assessment in childhood and adolescence. While studies have found evidence supporting different symptom clusters of disruptive behaviors in boys, research examining symptom clusters in girls is rare (Frick et al., 1993) and even rarer for girls in middle childhood, defined as ages 5 to 11 (e.g., Vaillancourt et al., 2003). Particularly wanting is research that empirically identifies distinct disruptive behavior symptom clusters in girls during middle childhood that (a) are based on a broad assessment of symptoms of oppositional behavior/conduct problems, inattention, hyperactivity/impulsivity, relational aggression, and callous-unemotional behavior; and (b) examines the relative stability of these behaviors across the developmental period of middle childhood. The present paper fills this gap by presenting exploratory and confirmatory factor analytic data on four waves of parent and teacher ratings of disruptive behavior problems in a large longitudinal study of girls during middle childhood.

---

Historically, a distinction has been made between ODD, CD, and ADHD symptoms. Although most of the research on the distinctiveness of these symptoms has been based on boys, the Diagnostic and Statistical Manual (DSM-IV; American Psychiatric Association (APA), 1994; Frick et al., 1991) clearly assumes that the distinct syndromes also apply to girls. However, research also suggests that the predominantly inattentive type of ADHD seems particularly relevant for girls (APA, 1994; Faraone, Biederman, Weber & Russell, 1998; Ford, Goodman & Meltzer, 2003). Some factor analytic studies have focused on ODD- and CD-symptoms. For example, Storvoll et al. (2002), in their study on boys and girls, found evidence for a rule-breaking factor (called nondestructive covert), an overt factor (verbal and nonverbal aggression), and covert factor (called destructive covert). This study reported that a three-factor solution fit the data for girls better than for boys. The study did not, however, include measurements of ADHD-symptoms. Other factor analytic studies have produced distinctive factors pertaining to ODD, CD, and ADHD (Hartman et al., 2001), but cross-loadings of symptoms across factors suggests that certain behaviors may not adequately distinguish between different factors (Hartman et al., 2001). Not all of these studies have carefully examined the extent to which factors based on large samples applied equally well across gender (e.g., Hartman et al., 2001).

Another line of research has focused on whether the concept of aggression should be expanded to include indirect or relational forms of aggression (e.g., making prank phone calls, writing critical notes or e-mails about another person; Crick & Grotpeter, 2005). There is evidence that girls who engage in these behaviors are at risk for negative social and emotional outcomes and that relational aggression adds unique variance in predicting these outcomes above and beyond physical aggression (Crick, Ostrov, & Werner, 2006). It is unclear, however, whether indirect aggression loads on either ODD or CD clusters of symptoms in girls.

Another set of studies has focused on features that seem to be early manifestations of adult psychopathy, such as interpersonally callous behaviors (Frick, O'Brien, Wootton, & McBurnett, 1994; Pardini, Obradović, & Loeber, 2006). While definitions vary, most of the existing research has focused on the identification of the affective features of adult psychopathy such as a lack of remorse or guilt, a lack of empathy, and shallow emotions (Frick et al., 1994). Studies have found that callous and unemotional behaviors can be reliably assessed and distinguished from traditional conceptualizations of CD, ODD, and ADHD in child and adolescent mixed samples of boys and girls (Dadds, Fraser, Frost, & Hawes, 2005; Frick, Bodin, & Barry, 2000) and in boys (Pardini et al., 2006).

In summary, although factor analytic studies have included girls, we have been unable to document factor analytic studies that have examined a broad array of disruptive behaviors, including the symptoms of ODD, CD, and ADHD, and items representing relational aggression and callous/unemotional behavior. We believe that only when all of these behaviors are included in factor analyses, is it possible to discern whether models of a single externalizing factor or two or more factors best fit the data. There are at least two reasons why middle childhood may be an important period for the emergence of girls' disruptive behaviors. First, a small proportion of girls develop these behaviors during those years. Second, more typical is that girls develop disruptive behavior during adolescence, but we know next to nothing about what might be precursors during middle childhood of adolescent disruptive behavior. Also, it remains to be seen whether externalizing factors are stable in girls during middle childhood.

The presence of different factors of externalizing behaviors in girls may depend on whether parents or teachers are the informants (Achenbach, McConaughy & Howell, 1987; De Los Reyes & Kazdin, 2005; Pulkkinen, Kaprio, & Rose, 1999). Each type of rater tends to have different opportunities for observation and different personal perspectives from which to assess

child disruptive behavior. Given these differences, little is known about the ways in which dimensions of externalizing behavior vary by informant.

In summary, the paper addresses the following questions:

1. Using exploratory factor analyses, what is the factor structure of girls' disruptive behavior between ages 5 and 11, and does the factor structure vary depending on whether parent or teacher ratings are used?
2. Does confirmatory factor analysis indicate that this factor structure fits the observed data well across the four different age cohorts over a period of four years?
3. How stable are the factors during middle childhood?

The questions are addressed in a large sample of Caucasian and African-American inner-city girls, who were followed up four times with little attrition. Because exploratory factor analyses can produce spurious results, confirmatory factor analyses were also used on the four years of data.

## Methods

### Sample Description

The participants of the Pittsburgh Girls Study (PGS) are 2,451 five- to eight-year-old girls recruited from a sample of 103,238 households in the city of Pittsburgh. Participants were identified by a stratified sampling of households in Pittsburgh neighborhoods; households in low-income neighborhoods were over-sampled. For the purposes of this study, neighborhoods were deemed low-income if at least 25% of the families were living at or below the poverty level, using 1990 Census data on poverty. Enumeration was completed in 89 of the 90 City of Pittsburgh neighborhoods during 1999, when households in low-income neighborhoods were fully enumerated (i.e., all households were contacted to determine eligibility for the study) while half of the households in other neighborhoods were randomly sampled. In total, 3,241 girls in the 5- to 8-year old age range – 83.7% of the girls noted in the 2000 Census – were identified. Of those girls initially identified as meeting the age criterion, 2,876 were asked to take part in the longitudinal study. From this pool, a total of 2,451 (85.2%) girls agreed to participate (for further details, see Hipwell, Loeber, Stouthamer-Loeber et al., 2002).

At the time of the first interview, the sample comprised 588 five-year-olds, 630 six-year-olds, 611 seven-year-olds, and 622 eight-year-olds. African American girls made up slightly more than half of the sample (52.8%), while 40.9% were Caucasian. Most of the remaining 6.3% of girls were described by their parents as multi-racial. In 92.7% of the interviews, the primary caregiver was a biological parent and in 92.9% of the cases the interviewed caregiver was female. To avoid using lengthy terminology, the word parent will be used to refer to the primary caregiver. Nearly all of the parents (83.2%) had at least a high school education. In a majority of households (58.78%), the parent was cohabiting with a spouse or domestic partner. Of the families surveyed, 38.9% reported receiving public assistance in the form of the Women, Infants, and Children supplemental nutrition program (WIC), food stamps, or welfare.

### Data Collection

Human subjects review and appropriate parental consent and child assent were obtained. Separate in-home interviews for both the child and the parent were conducted annually by trained interviewers using a laptop computer. Parents provided additional information by completing and returning a booklet of questionnaires. Teacher participation was obtained using questionnaire booklets, distributed via a mix of mail and hand-delivery. All participants were compensated for their involvement.

This paper covers the first four waves of parent and teacher data collected by the PGS. During this period of time, cohort 5 girls ranged in age from 5 to 8 (average age 5.6), cohort 6 were ages 6 to 9 (average age 6.7), cohort 7 were 7 to 10 years of age (average age 7.7), and cohort 8 ranged from 8 to 11 years old (average age 8.8). Because the girls were not interviewed at age 5 and a full interview was not administered until age 7, self-reported data from the girls were not used.

As noted previously, the PGS began its initial wave of data collection with 2,451 girls. All parents completed the interview during the first year. Valid teacher booklets were obtained from 1,832 (74.8%) of the participants' teachers during this wave. In year 2, interviews were completed by 2,383 (97.2%) parents, while 2,145 (87.5%) teachers completed and returned booklets. Parent participation was 95.4% (2,339 out of 2,451) and teacher participation was 84.8% (2,079 / 2,451) for the third interview year. In year 4, parent and teacher participation rates were 94.3% (2,310 out of 2,451) and 83.8% (2,054 / 2,451), respectively.

To assess the uniformity of the data across informants at each time point, attrition analyses were run. For each of years 2 through 4 (all parent interviews were completed in year 1), participants who had missing parent data were compared to those who completed the survey on girls' race (African American, Caucasian, Other), single parent status, household public assistance, and low parental education (parent with less than 12 years of formal education). A similar attrition analysis was also completed for years 1 through 4 of the teacher interview.

Among parents, year 2 difference in 'missingness' were found for girls living in single-parent households compared to those who were not (1.7% vs. 3.6%;  $\chi^2(1) = 7.43, p = 0.006$ ), as well as race (2.5% African American vs. 0% other race;  $\chi^2(1) = 4.26, p = 0.039$ ; 3.7% Caucasian vs. 0% other race;  $\chi^2(1) = 6.36, p = 0.012$ ). The analysis of year 3 parent data yielded only one significant result: differences were detected for girls living in single-parent households vs. those not living in single-parent households (3.4% vs. 5.6%;  $\chi^2(1) = 6.00, p = 0.014$ ). No significant differences were detected in the year 4 data.

During year 1, the only difference was on the extent of missing teacher data by race: rates of 'missingness' were virtually identical for African American and Caucasian girls (21.7% and 18.2% respectively;  $\chi^2(1) = 3.21, p = 0.073$ ) while other minorities showed a significantly lower attrition rate (13.4%) than African Americans ( $\chi^2(1) = 4.69, p = 0.030$ ), but not Caucasians ( $\chi^2(1) = 1.82, p = 0.177$ ). In year 2, girls with missing teacher data differed on receipt of public assistance (14.5% of girls whose family received public assistance versus 10.4% of girls whose family did not;  $\chi^2(1) = 8.77, p = 0.003$ ) and race – 14.8% of African Americans had missing data, which was significantly larger than both the percent of Caucasians with missing data (9.9%;  $\chi^2(1) = 12.76, p < 0.001$ ) and other races with missing data (6.7%;  $\chi^2(1) = 8.08, p = 0.005$ ). For year 3, the only significant difference was on participants who lived in single-parent households compared to those not living in single-parent households (12.8% vs. 16.5%;  $\chi^2(1) = 6.19, p = 0.013$ ). The analysis of year 4 data again showed a significant difference in the rate of missing teacher data by race, but only between African Americans and Caucasians (17.9% vs. 14.1%;  $\chi^2(1) = 6.18, p = 0.013$ ).

In summary, most of the attrition tests were nonsignificant. In the case of the parents, attrition was slightly more in the direction of two-parent compared to one-parent families (but the percentage difference was small), while in the case of the teachers, attrition was more in the direction African American compared to other families, but again not for all years.

## Measures

**Children's Peer Relationship Scale (CPRS, Crick & Grotpeter, 1995; Crick, 1996)**—The CPRS measures child-peer relations through frequencies of behaviors that reflect

six different domains of social functioning: perceived peer acceptance, isolation from peers, negative affect, engagement in caring acts, engagement in overt aggression, and engagement in relational aggression. The relational aggression subscale was comprised of items such as: ‘When some kids are mad at someone, they get back at the person by not letting the person in their group anymore’ scored on a 5-point (1–5) answer format, ranging from *never* to *almost always*. Crick and Grotpeter (1995) reported internal consistency of .73 on the relational aggression subscale (5 items). They also reported support for construct validity: Relationally aggressive girls were more disliked and reported poorer acceptance by peers than did nonaggressive girls, nonaggressive boys and relationally aggressive boys. Evidence of concurrent validity has been reported by Leadbeater, Banister, Ellis and Yeung (in press), who found that adolescent reports of relational aggression against peers were positively associated with relational dating aggression. The PGS administered adapted versions of the relational aggression subscale to the parent (5 items) and teacher (7 items). Two items were included in the teacher ratings that could be better observed in the school than in the home setting: ‘tells lies about peers’ and ‘ignores children or stops talking to him/her’.

**Child Symptom Inventory-4 (CSI-4, Gadow & Sprafkin, 1994)**—Items assessed the nature and severity of childhood behavioral disorder symptoms using criteria found in the Diagnostic and Statistical Manual of Mental Disorders – Fourth Edition (DSM-IV), including Conduct Disorder (CD), Oppositional Defiant Disorder (ODD) and Attention Deficit Hyperactivity Disorder (ADHD). These subscales have shown good sensitivity and specificity (using both parent and teacher reports) in distinguishing youth with clinical diagnoses from healthy controls (Gadow & Sprafkin, 1994). Each symptom was scored on 4-point (0–3) scales of *never*, *sometimes*, *a lot*, and *all the time*. Parent and teacher interviews each included 9 inattention items, 9 hyperactivity-impulsivity items, and 8 ODD items; however the parent interview included 13 CD items, while teachers were administered only 5 items. In the first year of data collection, symptoms were assessed for lifetime occurrence. In all ensuing years, only past year occurrence was assessed.

**Antisocial Processes Screening Device (APSD, Frick & Hare, 2001; Frick et al., 2000)**—The parent and teacher report of APSD was used to assess behaviors characteristic of a callous and unemotional (CU) interpersonal style. The CU subscale has shown good predictive validity in previous studies by distinguishing a group of children with conduct problems who develop particularly severe and aggressive behavior problems (Frick et al., 2005). The dimension consists of the following items: (a) ‘Is concerned about the feelings of others’ (reverse-scored); (b) ‘Feels bad or guilty when she does something wrong’ (reverse-scored); (c) ‘Is concerned about doing well in school’ (reverse-scored); (d) ‘Does not show feelings or emotions’ (e) ‘Is good at keeping promises’ (reverse-scored), and (f) ‘Keeps the same friends’ (reverse-scored). All six items were administered to the parent, while only four items (a – d) were administered to the teacher. The items were scored on a 3-point scale (0–2: *definitely true*, *sometimes true*, and *not at all true*, respectively) for behaviors occurring during the previous 2-month period for the teacher, while the reference period for the parent is the past year.

## Data Analysis

For this paper, parent and teacher report measures were examined separately using the same analytic strategy (for rationale, see Hartman et al., 1999). We were confronted with two possible approaches: combine parent- and teacher-information and do one factor analysis, or keep the informants separate and execute two separate factor analyses. The main disadvantages of the first option are that each informant observes different types of behaviors in different settings and that different factor structures might apply to each. For this reason, we chose the second option.



Although we aimed to include the same items for the analyses with parent and teacher, this was not always possible because teachers were unlikely to observe certain problem behaviors (e.g., curfew violation) and could therefore not provide reliable reports. Also, for reasons of economy of administration a few disruptive items were measured in the self-reported delinquency questionnaire (Loeber, Farrington, Stouthamer-Loeber, & Van Kammen, 1998), but because of scaling differences could not be included in the present analyses. For these reasons, the number of items included in analyses for parents differed occasionally from the number of items included in analyses for teachers. Separate analyses were undertaken to address whether these differences mattered for the results of the factor analyses.

Teacher ratings were unavailable in year 1 for all 5-year-olds, because most girls of that age were not yet participating in any type of formal schooling. Due to this lack of complete information, all teacher data on 5-year-old girls was dropped from the year 1 analysis. Teacher ratings of 6- to 8-year old girls were obtained for 74.75% of the girls during year 1 data collection. Analysis of all subsequent waves of teacher data included all available information. Additionally, due to the sampling technique used at recruitment (i.e. oversampling of girls in low income neighborhoods), a weighting variable was applied to all analyses in order to obtain parameter estimates for the general population of girls in Pittsburgh (see details in Hipwell et al., 2002).

A cursory examination of the prevalence rate for each item showed that the majority of the CD items had very few affirmative responses (i.e., less than 5% combining responses: *sometimes*, *a lot*, and *all the time*) at each wave of data collection. Preliminary exploratory factor analyses showed that these low base rate items were problematic in that they either did not load, or loaded only, on factors that did not make theoretical sense. To rectify this problem, conceptually similar items were combined together to form *testlets*. Specifically, conceptually similar low base rate items were combined into a single binary indicator (i.e., testlet) that was coded “1” if the girl had engaged in any of the behaviors assessed and “0” if she had refrained from engaging in all behaviors. For parent data, this entailed combining ‘used a weapon when fighting’, ‘forced sexual activity’, and ‘stolen things using physical force’ into a *serious violence* testlet; combining ‘cruel to animals’ and ‘cruel to people’ into a *cruelty* testlet; combining ‘vandalism’ and ‘deliberately started fires’ into a *destroy* testlet; and combining ‘broken into house, building, or car’ and ‘stolen when others were not looking’ into a *steal* testlet. The remaining questions – ‘bullied, threatened, or intimidated’, ‘started physical fights’, ‘lied to get things’ and ‘curfew violation’ – were left in the analyses as individual items because they had a higher prevalence. Because the teacher interview included only 5 CD items, a single teacher CD testlet was produced: *Serious Violence* (‘used a weapon when fighting’ and ‘stolen things using physical force’). Comparable to the analysis on the parent data, the remaining teacher items (‘bullied, threatened, or intimidated’, ‘cruel to people’, and ‘lied to get things’) remained in the analysis at an individual rather than at a testlet level.

Prior to examining the factor structure of the disruptive behaviors, the sample was randomly divided into two groups of about equal size: an experimental sample and a validation sample. The experimental sample comprised 1221 participants, while the validation sample included the remaining 1230 participants. Because the nature of the factor structure of disruptive behaviors is unknown, an exploratory factor analysis (EFA) was first performed on the experimental sample. The EFA utilized all of the available inattention, ODD, hyperactivity, impulsivity, relational aggression, and callous-unemotional items, plus the CD testlets and remaining individual items as described previously. The EFA analyses were run using a mean and variance-adjusted weighted least squares estimator (WLSMV) with an oblique rotation (promax) that allowed for correlated factors in Mplus 4.0 (Muthén & Muthén, 2006). There is evidence that the WLSMV estimator is optimal for estimating factor analysis parameters with ordinal data (Flora & Curran, 2004). With the EFA, all items are allowed to load on all factors

with the number of factors extracted being determined using several criteria including the examination of scree plots of successive eigenvalues, parallel analysis (O'Connor, 2002), and evaluating the interpretability of the solution. The procedure for the parallel analysis involved generating 1000 pseudo-random data sets (see O'Connor, 2002) for each year that contained the same number of observations as the number of completed interviews for that year (i.e., for year 1 teacher data, all 1000 generated data sets had 1832 observations). Next, exploratory factor analysis was run on each generated data set in a given year in order to generate the average eigenvalue expected for each successive factor extracted if the observations were randomly generated. These randomly generated eigenvalues were compared to the eigenvalues computed from the factor analysis of the PGS data to determine the maximum number of factors that could be reasonably extracted from the data. Specifically, only the factors corresponding to actual eigenvalues that were greater than the average eigenvalues generated from the parallel analysis were retained. However, additional factors were only retained if they consisted of a coherent group of common items that seemed to uniquely identify a construct.

After the appropriate number of factors was determined, the factor loading pattern matrix was examined to identify whether individual items exhibited evidence of consistent loadings on a single factor across multiple years. Specifically, the strength of item loadings was considered poor if they did not reach a value of .35 in at least three of the four years examined. Also, items were considered to discriminate too poorly between factors if the items exhibited loadings greater than or equal to .35 on more than one factor across two or more years.

Items from the EFA that were found to consistently load on a single factor across time were then submitted to a CFA using the validation sample of girls. With CFA, the number of factors must be specified *a priori* – a particular factor structure must be specified where items are forced to load on a particular factor.

The CFA analyses were also conducted using mean and variance-adjusted weighted least squares estimator (WLSMV) in Mplus 4.0 (Muthén & Muthén, 2006). We assessed absolute fit of the confirmatory models using global fit indices, including the comparative fit index (CFI), the Tucker-Lewis index (TLI), and the root mean square error of approximation (RMSEA). For the CFI and TLI, we used the conventional definition of values between .90 and .94 for acceptable fit, and .95 or greater for good fit. RMSEA values between .05 and .08 represent an acceptable fit, while values less than .05 indicate a good fit (McDonald & Ho, 2002).

## Results

### Exploratory Factor Structure Using Parent and Teacher Information

An exploratory factor analysis of the 45 parent items yielded nine eigenvalues greater than 1.0 in the first year and eight eigenvalues greater than 1.0 in each of the remaining three years. For the 41 teacher items, four eigenvalues were greater than 1.0 in year 1, while five eigenvalues exceeded the threshold in each of year 2 through year 4. As a further means of determining the factor structure, scree plots for both informants at each wave of data were examined to determine the point at which the plotted eigenvalues noticeably changed direction. These plots suggested that a five-factor solution for both parent and teacher informants was the most suitable choice. Parallel analysis suggested a maximum of eight factors could have been extracted from the year 1 parent data, with a limit of six factors for years 2–4. Similarly, the teacher data could have yielded up to four factors in years 1, 2, and 4, but a maximum of five factors was allowable in year 3.

Because interpretability was an important concern, both five- and six-factor rotated solutions were evaluated for the parent data. While the five-factor solution consistently fashioned the

same general conceptually valid factor structure (see Table 1), the six-factor solution never yielded a meaningful 'extra' factor. The resulting five parent factors were: (1) Inattention, (2) ODD/CD, (3) Hyperactivity-Impulsivity, (4) Relational Aggression, and (5) Callous-Unemotional Behavior.

The teacher data, however, were a little more troublesome since parallel analysis showed that a five-factor solution was not plausible at most time-points. Therefore, a four-factor solution was considered. The resulting four teacher factors were: (1) Inattention, (2) ODD/CD/Callous Unemotional Behavior, (3) Hyperactivity-Impulsivity, and (4) Relational Aggression.

## Parents

Table 2 shows the range of parent factor loadings for each item across the first four waves of data, plus the number of times each item had a 'significant' loading ( $\geq 0.35$ , as previously defined). Inattention was the most stable factor as it was the first factor extracted in all four years. Nine items loaded on this factor: 'failed to give close attention to details or made careless mistakes', 'had difficulty paying attention to tasks or play activities', 'seemed not to listen when spoken to directly', 'had difficulty following through on instructions and failed to finish things', 'had difficulty organizing tasks and activities', 'avoided doing tasks that require a lot of mental effort like schoolwork', 'lost things necessary for activities', 'easily distracted by other things going on', and 'forgetful in her daily activities'. Each of the inattention items loaded significantly on this factor at every time-point and none of these items cross-loaded on any other factor.

The second factor extracted at two of the four time-points dealt with Callous-Unemotional traits. Four items ('is concerned about how well she does at school or while doing tasks or activities', 'is good at keeping promises', 'feels bad or guilty when she does something wrong', and 'is concerned about the feelings of others') significantly loaded for all four years, while two additional items ('does not show feelings or emotions' and 'keeps the same friends') loaded in three of the four years. None of these items showed any tendencies to cross-load.

Items associated with ODD and CD loaded on a conduct problems dimension – the third factor extracted in all four years. The 12 items meeting criteria for inclusion in this factor were: 'lost her temper', 'argued with adults', 'defied you or refused to do what you told her to do', 'done things to deliberately annoy others', 'blamed others for her own misbehavior or mistakes', 'touchy or easily annoyed with others', 'angry and resentful', 'taken her anger out on others or tried to get even', 'bullied, threatened or intimidated others', 'started physical fights', the serious violence testlet, and the cruelty testlet. All other CD and ODD items were subsequently removed from future analysis as they did not meet the threshold for inclusion.

The fourth factor extracted in three of the four years consisted of seven hyperactivity-impulsivity items. Items associated with this factor included: 'blurted out answers to questions before they have been completed', 'had difficulty waiting her turn in group activities', 'interrupted people or butted into other children's activities', 'run about or climbed on things when asked not to do so', 'had difficulty playing quietly', 'acted as if driven by a motor or on the go', and 'talked excessively'. Two items – 'fidgeted with her hands or feet or squirmed in her seat' and 'had difficulty remaining seated when asked to do so' – were removed because they loaded on more than one factor.

The final dimension extracted concerned characteristics of relational aggression. The resulting factor was comprised of five items: 'excludes others to get even', 'spreads rumors', 'tries to get other children to stop playing/liking him/her', 'threatens to stop being friend' and 'excludes others from peer group activities'. All five items loaded on this factor for each of the four years with no cross-loading on other factors.



## Teachers

The range of teacher factor loadings for years 1–4, along with the number of times each item significantly loaded on a factor, are shown in Table 2. As with the parent analysis, the inattention factor was the first one extracted in all four years. The structure of the teacher factor is identical to that of the parent version, including the same nine items.

The second factor found in three of the four years was the hyperactivity-impulsivity factor. This dimension differed from the parent analogue in that only one item ('fidgeted with her hands or feet or squirmed in her seat') was removed due to extreme cross-loading. Items associated with the teacher hyperactivity-impulsivity factor included: 'blurted out answers to questions before they have been completed', 'had difficulty waiting her turn in group activities', 'interrupted people or butted into other children's activities', 'had difficulty remaining seated when asked to do so', 'ran about or climbed on things when asked not to do so', 'had difficulty playing quietly', 'acted as if driven by a motor or on the go', and 'talked excessively'.

In all four years, the combined ODD/CD and callous-unemotional behavior factor was extracted third. This factor was comprised of eight ODD items, three CD items, and the only two callous-unemotional items which loaded according to the rules previously set forth. The items which loaded on this factor were: 'lost her temper', 'argued with adults', 'defied or refused to do what she was told', 'done things to deliberately to annoy others', 'blamed others for her own misbehavior or mistakes', 'touchy or easily annoyed with others', 'angry and resentful', 'taken her anger out on others or tried to get even', 'cruel to people', 'lied to get things', the serious violence testlet, 'feels bad or guilty when she does something wrong' and 'is concerned about the feelings of others' (reverse coded). Of the other items that consistently load on this factor, one – 'is concerned about how well she does at school or while doing tasks or activities' – cross-loads with the inattention factor at all four time-points, while the other – 'bullied, threatened or intimidated others' – cross-loads with the relational aggression factor each year. Both of these items were removed from further analysis.

The final teacher factor – extracted last in three of the four years – was relational aggression. Seven items met the criteria for inclusion into this factor: 'excludes to get even', 'spreads rumors', 'tries to get other children to stop playing/liking him/her', 'tells lies about peers', 'threatens to stop being friend', 'ignores child or stops talking to him/her', and 'excludes from peer group activities'.

## Confirmatory Factor Analysis

A confirmatory factor analysis using the WLSMV estimator was modeled on parent data using the validation sample. Ranges for the estimated CFA standardized factor loadings are presented in Table 3. Each of the ODD/CD, inattention, hyperactivity-impulsivity, and relational aggression items consistently loaded above the .35 threshold across all four years. Four of the six callous-unemotional items loaded above the cut-off point at all four time-points, but 'is concerned about how well she does at school or while doing tasks or activities' did not achieve the threshold in any of the four years, while 'does not show feelings or emotions' missed the cut-point in two years.

Teacher CFA results are presented in Table 3. Standardized factor loadings for the model were fairly large ( $> .54$ ) for each of the four teacher-reported factors, with the majority of loadings larger than .80. In general, the items with the smallest loadings are the two callous-unemotional items that loaded with ODD/CD in the EFA; however these two items still have relatively large loadings, ranging from .54 to .65 for 'feels bad or guilty when she does something wrong' and from .71 to .75 for 'is concerned about the feelings of others'.

Overall, items associated with factors derived from parent ratings were similar to items associated with equivalent factors derived from teacher ratings. For example, for the parent ODD/Conduct Problem factor, we found that items such as lost temper, argued with adults, defied/refused, bullied, and vindictive also loaded on the teacher version of this factor, except that some items associated with callous/unemotional behavior (i.e., 'feels bad when does something wrong', and 'concerned about feelings of others') also loaded onto this factor. Items that comprised the Inattention factors derived from parent ratings were identical to those included in the equivalent factor based on teacher ratings. The same applied to the Hyperactivity-Impulsivity factor derived from ratings by each informant (only 'difficulty remaining seated' was unique to the teacher factor). Finally, the overlap between items loading on the Relational Aggression factor based on parent ratings was very similar to that based on teacher ratings (the exceptions were 'threatens to stop being a friend', and 'excludes from group activities', which loaded on the teacher factor only). Thus, in general there was high consistency of items loading on equivalent factors derived from the two types of informants. The unique items in a few of the factors (e.g., 'difficulty remaining seated', and 'excludes others from group') based on teacher ratings may reflect teachers' superior ability to observe these behaviors in the school setting.

Table 4 lists the model fit statistics for each of the parent and teacher CFA models. Using the TLI, parent model fit was acceptable in year 1, but good for years 2–4. The RMSEA shows acceptable fit in all four years for the parent model. According to the CFI, though, the parent model did not fit well for the first three years, but the fit was acceptable in year 4. Conversely, the teacher model fit was acceptable in all four years based on the CFI, was good all four years based on the TLI, and fell just outside the acceptable range in all four years according to the RMSEA.

Because the items for parents and teachers differed somewhat, the factor analyses were repeated with only those items that were shared between the two informants. The results showed that the analyses produced the same factors for the parents and teachers, respectively, when based on the more restricted range of items.<sup>1</sup>

Correlations between the factors were also assessed and the results are reported in Table 5. For the parent model, results indicated that a moderate degree of intercorrelation existed between inattention, ODD/CD, and hyperactivity-impulsivity, with coefficients ranging from .51 to .66 ( $p < .001$ ). Relational aggression and ODD/CD were also moderately correlated ( $r$  ranging from .49 to .59,  $p < .001$ ). In general, the remaining parent factors all exhibited low to moderate intercorrelation ( $r$  ranging from .32 to .54,  $p < .001$ ). Among the teacher factors, inattention and relational aggression generally showed a low to moderate level of association ( $r$  from .39 to .54,  $p < .001$ ), but all other factors were more highly interrelated with correlations coefficients ranging from .57 to .84 ( $p < .001$ ).

### Temporal Stability of the Constructs

The last question we addressed was to what extent each of the identified factors was stable across time. To examine this we first constructed scale scores for each factor identified in the parent and teacher CFAs (see Table 6). Specifically, items that loaded on a factor were averaged together to create a composite score at each assessment wave. This was done separately for parent and teacher measures. Intraclass correlation coefficients (ICC) were then calculated using an absolute agreement specification between every one year interval for each parent and teacher construct. In addition, an average ICC was calculated across all four waves for each parent and teacher construct to estimate temporal stability across all time points. For these

---

<sup>1</sup>The results can be requested from the senior author.

analyses, the ICC is the proportion of a variance in the construct that is consistent across different measurement occasions. Table 6 shows the year-to-year ICCs for each parent-derived factor for wave 1 to 4. All ICCs were statistically significant ( $p < .001$ ), and the year-to-year stability estimates were high for all constructs (all ICCs  $\geq .71$ ). For each factor the year-to-year ICCs increased slightly from Year 1 to Year 4.

Teacher intraclass correlations are reported in Table 6. All correlations were statistically significant ( $p < .001$ ), and were moderate to high. The least stable construct was the relational aggression construct, which had an overall ICC of .69 across years 1–4. The stability estimates for the remaining constructs were similar to those found for the parent-report constructs.

## Discussion

This study examined a broad range of young girls' disruptive behaviors (oppositional behavior/conduct problems, inattention, hyperactivity/impulsivity, relational aggression, and callous-unemotional behavior) as rated by parents and teachers over a period of four years. On the basis of exploratory and confirmatory factor analyses, the results showed that the factor structure of disruptive behavior best fit five factors for parent ratings and four factors for teacher ratings with largely similar item contents across different informants. These results agree with findings by Pulkkinen, Kaprio, and Rose (1999) showing that parent ratings result in a more differentiating view of individual child differences than teacher ratings. We found that three factors were mainly similar for parents and teachers: inattention, hyperactivity-impulsivity, and relational aggression. The majority of items that loaded onto factors derived from parent ratings were identical to items loading on equivalent factors derived from teacher ratings. The parent-rated conduct/oppositional factor was distinct from the callous-unemotional factor, but the two factors were combined when teachers were the informants. The intercorrelation between the parent-derived factors was moderately strong, while for the teacher-derived factors was slightly higher.

The stability from year to year of each of the factors was high for parents and slightly lower when teachers were the informants. These findings lend support to the notion that girls' clusters of disruptive behaviors represented by the factor analyses are quite stable in middle childhood, even when judged by different teachers in successive grades of elementary schools. We know from other studies that the stability of mental health problems in girls is either as high as in boys or higher (Costello, Mustillo, Erkanli, Keeler, & Angold, 2003; Tremblay et al., 1992; Verhulst & Vander Ende, 1991; Zoccolillo, Pickles, Quinton, & Rutter, 1992). However, to our knowledge, information on the stability of specific clusters of disruptive behaviors in girls is not available from other studies.

How do the results of the factor analyses compare with other studies? We have been unable to find a study on the factor structure of girls' disruptive behavior with repeated measurements between ages 5 and 11. However, single-wave factor analytic studies have shown distinctions between inattention and hyperactivity/impulsivity (e.g., Burns, Walsh, Owen, & Snell, 1997a; Burns et al., 1997b; Burns, Boe, Walsh, Somers-Flanagan, & Teegarden, 2001; Burns & Patterson, 2000), and between indirect and physical aggression (Vaillancourt et al., 2003). Some studies have found a distinction between ODD- and CD-symptoms, but this has been for populations that are older than in the present study (Burns et al., 1997). It is not uncommon that CD-symptoms, because of their low prevalence, are omitted from factor analyses (e.g., Burns et al., 2001). Only a few studies have examined the factor structure for girls, and they have not presented detailed results other than stating that results applied to each gender (e.g., Burns et al., 1997b). Finally, not all studies have demonstrated subclusters of symptoms within the externalizing syndrome (Hartman et al., 1999).

There can be several reasons why a five-factor solution was found for parents and a four-factor solution for teachers. As mentioned, the parent-rated conduct/oppositional factor was distinct from the callous-unemotional factor, but the two factors were combined when teachers were the informants. One possibility is that different child behaviors are manifest in the school compared to home and that the callous-unemotional behaviors either are less common (or less likely to be observed by teachers) in school or are more often accompanied by conduct/oppositional behaviors in school than in the home setting. Another possibility is that the changing of teachers each year, compared to the presence of the same parents as raters of the four years, put parents more at an advantage in observing a distinct pattern of callous/unemotional behavior. Perhaps the most important issue is the predictive validity of later psychopathology in the girls when based on the respective parent- and teacher-observed factors, and whether this differs among the two types of informants in the long term. Only with long-term information to hand will we be able to say whether routine assessments of girls by means of parent and teacher ratings should be based on the observed factor scores. The continued follow-up of the girls into adolescence will clarify this matter.

How do the present results for girls compare with factor analytic studies on boys? Cross-sectional studies on boys show replicated findings on three to four distinct factors of oppositional defiant behavior toward adults, conduct problem behavior, inattentive behavior, and hyperactivity-impulsivity (e.g., Burns & Patterson, 2000; Lahey et al., 2008). Other studies have found a distinction between overt (aggressive) and covert conduct problems (e.g., Frick et al., 1991; Loeber & Schmalzing, 1985). Most factor analyses conducted on data concerning boys have not included measurements of callous/unemotional behaviors. Among the exceptions are Pardini, Obradović and Loeber (2006), who found that a four factor solution of interpersonal callousness, hyperactivity/impulsivity, inattention, and conduct problems best fit the data (see also Dadds et al., 2005). However, we have been unable to find studies that also measured relational aggression. Thus, it is unclear whether relational aggression in boys constitutes an independent factor once other disruptive behaviors are measured as well.

The present study supports the notion that consideration of clusters of symptoms of disruptive behavior is preferable to consideration of all types of disruptive behavior as a single cluster. Many of the symptoms of oppositional behavior/conduct problems, inattention and hyperactivity/impulsivity ODD, CD, are routinely included in assessment instruments. However, on the basis of the current findings, we recommend that future measurement instruments should include items representative of relational aggression and of callous-unemotional behavior, respectively. Currently, DSM-IV (APA, 1994) does not contain measurements for these dimensions, and we hope that plans for DSM-V will include a discussion of the utility of symptoms representative of relational aggression and callous-unemotional behaviors for assessments in girls and boys (for boys, see Pardini et al., 2006).

The study has some limitations. The number and type of measurement items per factor varied and were not exactly the same for each informant. Not all items from the original measures were administered to the informants including items from the teacher-reported APSD assessing CU features in girls. The study focused on girls rather than boys and does not report on similarities and differences in the factor structure across the two genders. Invariably, disruptive behaviors with a low frequency between ages 5 and 11 had to be removed from the analyses, and it remains to be seen how the factor structure will be at later ages when such behaviors will increase in populations of girls. Given that cross-informant factor analyses often produce poor results (Hartman et al., 1999), we analyzed data separately for parent and teacher. Against these limitations, the study has several strengths such as yearly, repeated ratings by parents and teachers from middle to late childhood, completeness of data, representativeness and large size of the sample, and the examination of the stability of factor structure of disruptive behavior with age.

We are continuing to follow up the girls in the Pittsburgh Girls Study and we anticipate that as the girls enter adolescence the types, diversity, and severity of their disruptive behavior will change. It is an open question whether the dimensional structure of disruptive behavior in girls will change with development as well.

## Acknowledgments

We are much indebted to the staff of the Pittsburgh Girls Study for their dedication and hard work in conducting the study. The study was financed by grants from the National Institute of Mental Health (MH 56630) and the National Institute on Drug Abuse (DA012237).

## References

- Achenbach TM, McConaughy SH, Howell CT. Child/adolescent behavioral and emotional problems: Implications of cross-informant correlations for situational specificity. *Psychological Bulletin* 1987;101:213–232. [PubMed: 3562706]
- American Psychiatric Association. DSM-IV. Vol. fourth edition. Washington, D.C.: American Psychiatric Association; 1994.
- Burns GL, Patterson DR. Factor structure of the Eyberg Child Behavior Inventory: A parent rating scale of oppositional defiant behavior toward adults, inattentive behavior, and conduct problem behavior. *Journal of Clinical Child Psychology* 2000;29:569–577. [PubMed: 11126634]
- Burns GL, Walsh JA, Owen SM, Snell J. Internal validity of Attention Deficit Hyperactivity Disorder, Oppositional Defiant Disorder, and Overt Conduct Disorder symptoms in young children: Implications from teacher ratings for a dimensional approach to symptom validity. *Journal of Clinical Child Psychology* 1997a;26:266–275. [PubMed: 9292384]
- Burns GL, Walsh JA, Patterson DR, Holte CS, Sommers-Flanagan R, Parker CM. Internal validity of the Disruptive Behavior Disorder Symptoms: Implications from parent ratings for a dimensional approach to symptom validity. *Journal of Abnormal Child Psychology* 1997b;25:307–319. [PubMed: 9304447]
- Burns GL, Boe B, Walsh JA, Somers-Flanagan R, Teegarden LA. A Confirmatory factor analysis on the DSM-IV ADHD and ODD symptoms: What is the best model for the organization of these symptoms? *Behavioral Science* 2001;29:339–349.
- Costello EJ, Mustillo S, Erkanli A, Keeler G, Angold A. Prevalence and development of psychiatric disorders in childhood and adolescence. *Archives of General Psychiatry* 2003;60:837–844. [PubMed: 12912767]
- Crick NR. The role of overt aggression, relational aggression, and prosocial behavior in the prediction children's future social adjustment. *Child Development* 1996;67:2317–2327. [PubMed: 9022243]
- Crick NR, Grotpeter JK. Relational aggression, gender, and social-psychological adjustment. *Child Development* 1995;66:710–722. [PubMed: 7789197]
- Crick NR, Ostrov JM, Werner ME. A longitudinal study of relational aggression, physical aggression, and children's social-psychological adjustment. *Journal of Abnormal Child Psychology* 2006;34:127–138.
- Dadds MR, Fraser JA, Frost AD, Hawes DJ. Disentangling the underlying dimensions of psychopathy and conduct problems in childhood: A community study. *Journal of Consulting and Clinical Psychology* 2005;73:400–410. [PubMed: 15982138]
- De Los Reyes A, Kazdin AE. Informant discrepancies in the assessment of childhood psychopathology: A critical review, theoretical framework, and recommendations for further study. *Psychological Bulletin* 2005;131:483–509. [PubMed: 16060799]
- Faraone SV, Biederman J, Weber W, Russell RL. Psychiatric, neuropsychological, and psychosocial features of DSM-IV subtypes of attention-deficit/hyperactivity disorder: Results from a clinically referred sample. *Journal of the American Academy of Child & Adolescent Psychiatry* 1998;37:185–193. [PubMed: 9473915]
- Flora DB, Curran PJ. An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with original data. *Psychological Methods* 2004;9:466–491. [PubMed: 15598100]



- Ford T, Goodman R, Meltzer H. The British Child and Adolescent Mental Health Survey 1999: The prevalence of DSM-IV Disorders. *Journal of the American Academy of Child & Adolescent Psychiatry* 2003;42:1203–1211. [PubMed: 14560170]
- Frick PJ, Bodin SD, Barry CT. Psychopathic traits and conduct problems in community and clinic-referred samples of children: Further development of the Psychopathy Screening Device. *Psychological Assessment* 2000;12:382–393. [PubMed: 11147105]
- Frick, PJ.; Hare, RD. *The antisocial process screening device*. Toronto: Multi-Health Systems; 2001.
- Frick PJ, Lahey BB, Loeber R, Stouthamer-Loeber M, Green S, Hart EL, Christ MAG. Oppositional defiant disorder and conduct disorder in boys: Patterns of behavioral covariation. *Journal of Clinical Child Psychology* 1991;20:202–208.
- Frick PJ, Lahey BB, Loeber R, Tannenbaum L, Van Horn Y, Christ MAG, Hart EA, Hanson K. Oppositional defiant disorder and conduct disorder: A meta-analytic review of factor analyses and cross-validation in a clinic sample. *Clinical Psychology Review* 1993;13:319–340.
- Frick PJ, O'Brien BS, Wootton JM, McBurnett K. Psychopathy and conduct problems in children. *Journal of Abnormal Psychology* 1994;103:700–707. [PubMed: 7822571]
- Frick P, Stickle T, Dandreaux D, Farrell J, Kimonis E. Callous-unemotional traits in predicting the severity and stability of conduct problems and delinquency. *Journal of Abnormal Child Psychology* 2005;33:471–487. [PubMed: 16118993]
- Gadow, KD.; Sprafkin, J. *Child Symptom Inventories Manual*. Stony Brook, NY: Checkmate Plus; 1994.
- Hartman CA, Hox J, Auerbach J, Erol N, Fonseca AC, Mellenbergh J, Novik TS, Oosterlaan J, Roussos AC, Shalev RS, Zilber N, Sergeant JA. Syndrome dimensions of the Child Behavior Checklist and the Teacher Report Form: A critical empirical evaluation. *The Journal of Child Psychology and Psychiatry and Allied Disciplines* 1999;40:1045–1116.
- Hartman CA, Hox J, Mellenbergh GJ, Boyle MH, Offord DR, Racine Y, McNamee J, Gadow KD, Sprafkin J, Kelly KL, Nolan EE, Tannock R, Schachar R, Schut H, Postma I, Drost R, Sergeant JA. DSM-IV internal construct validity; When a taxonomy meets data. *Journal of Child Psychology and Psychiatry* 2001;42:817–836. [PubMed: 11583254]
- Hipwell AE, Loeber R, Stouthamer-Loeber M, Keenan K, White HR, Kroneman L. Characteristics of girls with early onset disruptive and antisocial behaviour. *Criminal Behavior of Mental Health* 2002;12:99–118.
- Lahey BB, Rathouz PJ, Van Hulle C, Urbano RC, Krueger RF, Applegate B, Garriock HA, Chapman DA, Waldman ID. Testing structural models of DSM-IV symptoms of common forms of child and adolescent psychopathology. *Journal of Abnormal Child Psychology* 2008;36:187–206. [PubMed: 17912624]
- Leadbeater B, Banister E, Ellis W, Yeung R. Victimization and relational aggression in adolescent romantic relationships: The influence of parental and peer behaviors, and individual adjustment. *Journal of Youth and Adolescence*. (in press)
- Loeber, R.; Farrington, DP.; Stouthamer-Loeber, M.; Van Kammen, WB. *Antisocial Behavior and Mental Health Problems: Explanatory Factors in Childhood and Adolescence*. Mahwah, NJ: Lawrence Erlbaum; 1998.
- Loeber R, Schmalting K. Empirical evidence for overt and covert patterns of antisocial conduct problems. *Journal of Abnormal Child Psychology* 1985;13:337–352. [PubMed: 4008760]
- McDonald RP, Ho MH. Principles and practice in reporting structural equation analyses. *Psychological Methods* 2002;7:64–82. [PubMed: 11928891]
- Muthén, LK.; Muthén, R. *Mplus: The Comprehensive Modeling Program for Applied Researchers*. Los Angeles, CA: version 4; 2006.
- O'Connor BP. Comprehensiveness of the Five-Factor Model in relation to popular personality inventories. *Assessment* 2002;9:188–203. [PubMed: 12066834]
- Pardini D, Obradović J, Loeber R. Interpersonal callousness, hyperactivity/impulsivity, inattention and conduct problems as precursors to delinquency persistence boys: A comparison of three grade based cohorts. *Journal of Clinical Child and Adolescent Psychology* 2006;36:46–59. [PubMed: 16390302]
- Prinstein MJ, Boergers J, Vernberg EM. Overt and relational aggression in adolescents: social-psychological adjustment of aggressors and victims. *Journal of Clinical Child Psychology* 2001;30:479–491. [PubMed: 11708236]

- Pulkkinen L, Kaprio J, Rose RJ. Peers, teachers and parents as assessors of the behavioral and emotional problems of twins and their adjustment: The Multidimensional Peer Nomination Inventory. *Twin Research and Human Genetics* 2:274–285.
- Storvoll EE, Wichstrøm L, Kolstad A, Pape H. Structure of conduct problems in adolescence. *Scandinavian Journal of Psychology* 2002;43:81–91. [PubMed: 11885763]
- Tremblay RE, Masse B, Perron D, Leblanc M, Schwartzman AE, Ledingham JE. Early disruptive behavior, poor school achievement, delinquent behavior, and delinquent personality: longitudinal analyses. *Journal of Consulting and Clinical Psychology* 1992;60:64–72. [PubMed: 1556287]
- Vaillancourt T, Brendgen M, Boivin M, Tremblay RE. A longitudinal confirmatory factor analysis of indirect and physical aggression: Evidence of two factors over time? *Child Development* 2003;74:1628–1638. [PubMed: 14669886]
- Vaillancourt T, Miller JL, Fagbemi J, Côté S, Tremblay RE. Trajectories and predictors of indirect aggression: Results from a nationally representative longitudinal study of Canadian children aged 2–10. *Aggressive Behavior* 2007;33:314–326. [PubMed: 17593562]
- Verhulst FC, van der Ende J. Assessment of child psychopathology: relationships between different methods, different informants and clinical judgment of severity. *Acta Psychiatrica Scandinavica* 1991;84:155–159. [PubMed: 1950609]
- Zoccolillo M, Pickles A, Quinton D, Rutter M. The outcome of childhood conduct disorder: implications for defining adult personality disorder and conduct disorder. *Psychological Medicine* 1992;22:971–986. [PubMed: 1488492]

Table 1

Eigenvalues based on parent and teacher exploratory factor analyses

	Year 1	Year 2	Year 3	Year 4	Eigenvalue
Factor	Eigenvalue	Eigenvalue	Eigenvalue	Eigenvalue	
Parent					
Inattention	12.45	14.83	15.91	15.91	15.36
Relational Aggression	3.60	3.56	4.04	4.04	3.81
Conduct Problems	2.47	2.49	2.47	2.47	2.70
Hyperactivity- Impulsivity	2.06	2.25	2.00	2.00	2.39
Callous-Unemotional	1.87	2.00	1.95	1.95	1.96
Total	22.45	25.13	26.37	26.37	26.22
% of Variance	49.87	55.82	58.59	58.59	58.25
Teacher					
Inattention	24.13	23.39	23.04	23.04	23.79
Relational Aggression	4.80	4.76	4.62	4.62	4.58
Conduct Problems & Callous-Unemotional	2.13	2.22	2.33	2.33	2.32
Hyperactivity- Impulsivity	1.64	1.78	1.71	1.71	1.60
Total	31.70	32.15	31.70	31.70	32.29
% of Variance	79.76	78.41	77.32	77.32	78.76

Table 2

Items	P/T	ODD/CD & Callous Unemotional			Inattention			Relational Regression			Hyperactivity/Impulsivity			Callous-Unemotional		
		Min	Max	≥.35	Min	Max	≥.35	Min	Max	≥.35	Min	Max	≥.35	Min	Max	≥.35
Lost Temper	P	.81	.95	4	-.05	.06	0	-.07	-.03	0	-.13	.00	0	-.28	-.13	0
Argued w/ Adults	P	.77	.83	4	-.16	-.12	0	.03	.15	0	.15	.23	0	-.15	.23	0
Defied/Refused	P	.71	.86	4	-.16	-.08	0	-.01	-.12	0	.22	.24	0	-.28	-.14	0
Deliberately Annoys	P	.72	.79	4	.04	.09	0	-.26	-.15	0	-.03	.06	0	-.07	.08	0
Blames Others	P	.42	.57	4	-.05	.03	0	-.05	.03	0	.11	.07	0	.02	.15	0
Touchy/Easily Annoyed	P	.51	.63	4	.06	.15	0	.16	.28	0	.14	.22	0	.08	.14	0
Angry or Resentful	P	.46	.65	4	.00	.17	0	.09	.16	0	.18	.25	0	.08	.14	0
Vindictive	P	.59	.74	4	.02	.11	0	-.02	.10	0	.17	.22	0	-.17	.03	0
Bullied	P	.62	.82	4	-.01	.04	0	-.17	.22	0	.04	.16	0	-.11	.09	0
Started Physical Fights	P	.41	.58	4	-.05	.03	0	-.02	.13	0	-.16	-.07	0	.00	.21	0
Cruel to People	P	.65	.88	4	-.09	.01	0	-.10	.19	0	-.06	.05	0	.00	.22	0
Lied to Get Things	P	.71	.89	4	-.14	.01	0	.24	.33	0	.02	.14	0	.04	.22	0
Curfew Violation	P	.48	.64	4	-.26	.00	0	.21	.23	0	.06	.23	0	.06	.23	0
Serious Violence Testlet	P	.34	.54	4	-.10	.03	0	.35	.48	4	.16	.16	0	.04	.22	0
Cruelty Testlet	P	.43	.61	4	-.17	.03	0	.09	.17	0	-.08	.13	0	.06	.23	0
Destroy Testlet	P	.29	.48	2	-.03	.06	0	.08	.34	0	.03	.11	0	.13	.28	0
Steal Testlet	P	.12	.37	1	.08	.16	0	.03	.33	0	.05	.11	0	.21	.41	1
Careless Mistakes	P	-.12	.02	0	.64	.84	4	.09	.12	0	-.07	-.01	0	-.12	-.07	0
Difficulty Paying Attention	P	-.15	-.06	0	.82	.90	4	-.05	.12	0	.03	.08	0	-.06	.08	0
Not Listening	P	-.19	-.02	0	.71	.85	4	-.04	.04	0	.00	.12	0	.05	.17	0
Failed to Finish	P	-.09	-.02	0	.82	.93	4	-.03	.05	0	.07	.21	0	.00	.06	0
Disorganized	P	-.07	.05	0	.72	.85	4	-.06	.06	0	-.16	.03	0	-.08	.02	0
Avoided Mental Effort	P	-.10	.00	0	.93	1.00	4	-.05	.03	0	.00	.07	0	.00	.11	0
Lost Things	P	.04	.12	0	.68	.71	4	-.02	.02	0	-.19	-.05	0	.00	.04	0
Easily Distracted	P	.04	.13	0	.84	.89	4	.06	.07	0	-.08	-.06	0	-.13	.04	0
Forgetful	P	.02	.06	0	.40	.65	4	-.04	.14	0	.00	.08	0	.00	.07	0
Blurred Out Answers	P	-.12	.00	0	.76	.86	4	-.00	.06	0	.18	.24	0	-.15	.05	0
Difficulty Waiting Turn	P	-.11	-.06	0	.60	.73	4	-.03	.02	0	.16	.34	0	-.16	-.07	0
Interrupt Activities	P	-.08	-.01	0	.52	.78	4	-.00	.11	0	.07	.12	0	-.02	.02	0
	P	-.13	-.09	0	.89	.93	4	.08	.12	0	.56	.75	4	-.16	-.07	0
	P	-.07	.03	0	-.06	.01	0	-.03	.14	0	.80	.86	4	-.02	-.01	0
	P	-.04	.11	0	-.02	.19	0	.08	.17	0	.69	.76	4	-.10	.02	0
	P	.15	.26	0	-.11	.03	0	-.04	.16	0	.55	.69	4	-.02	.02	0
	P	.07	.15	0	-.07	.09	0	.10	.21	0	.10	.21	0	-.10	.02	0

Items	P/T	ODD/CD & Callous Unemotional			Inattention			Relational Regression			Hyperactivity/Impulsivity			Callous-Unemotional		
		Min	Max	≥.35	Min	Max	≥.35	Min	Max	≥.35	Min	Max	≥.35	Min	Max	≥.35
Fidgeted/Squirmed	T	.17	.24	0	-.12	.03	0	.14	.23	0	.65	.75	0	-.11	.02	0
	P	-.01	.05	0	.31	.44	3	-.12	.01	0	.33	.45	3			
	T	-.05	.13	0	.36	.45	4	-.16	-.06	0	.58	.67	4			
Difficulty Remaining Seated	P	.01	.03	0	.21	.42	2	-.22	-.11	0	.46	.62	4	.09	.24	0
	T	.02	.15	0	.27	.33	0	-.12	-.03	0	.63	.69	4			
	P	.03	.22	0	.07	.12	0	-.19	-.10	0	.43	.60	4	.10	.20	0
Run/Climb	T	.13	.26	0	.21	.38	1	-.12	-.06	0	.59	.67	4			
	P	-.03	.11	0	.10	.24	0	-.05	.02	0	.44	.65	4	.05	.15	0
Difficulty Playing Quietly	T	.07	.19	0	.14	.19	0	.00	.08	0	.59	.71	4			
'Driven by Motor'	P	-.09	.00	0	-.03	.07	0	-.09	-.07	0	.69	.79	4	-.14	-.06	0
	T	-.02	.06	0	.09	.14	0	-.12	-.01	0	.77	.80	4			
Talked Excessively	P	-.13	-.07	0	-.12	-.03	0	.01	.08	0	.73	.84	4	-.17	-.02	0
	T	-.06	.03	0	.00	.13	0	.12	.24	0	.70	.80	4			
Excludes to Get Even	P	.07	-.01	0	-.05	.00	0	.75	.87	4	-.01	.06	0	-.07	.07	0
	T	-.02	.15	0	-.09	-.01	0	.84	.93	4	.00	.06	0			
Spreads Rumors	P	.04	.11	0	-.07	.09	0	.65	.70	4	-.07	-.03	0	-.06	.02	0
	T	.06	.11	0	.02	.08	0	.81	.86	4	.00	.06	0			
Tries to Get Others to Stop Liking	P	-.10	-.02	0	-.02	.04	0	.88	.94	4	-.04	.04	0	-.02	.07	0
	T	.01	.06	0	.01	.04	0	.90	.92	4	-.02	.09	0			
Threatens to Stop Being Friend	P	-.03	.08	0	-.02	.02	0	.72	.82	4	-.02	.12	0	-.03	.09	0
	T	.00	.06	0	-.05	.00	0	.82	.84	4	-.05	.00	0			
Lies about Peers	T	.12	.17	0	.04	.16	0	.79	.83	4	-.04	.04	0			
	P	.04	.08	0	-.06	.05	0	.65	.77	4	-.06	-.03	0	-.11	-.06	0
Ignores	T	.03	.05	0	.06	.08	0	.82	.90	4	-.03	.07	0			
	P	-.22	-.11	0	-.07	.00	0	-.16	-.07	0	-.07	.07	0	.47	.53	4
Concerned About School work	T	.34	.43	3	.56	.67	4	-.18	-.01	0	-.26	-.14	0			
Excludes from Group Activities	T	.01	.08	0	.00	.06	0	.88	.92	4	-.02	.03	0			
Good at Keeping Promises	P	-.08	.06	0	.12	.22	0	.05	.06	0	-.07	.07	0	.47	.53	4
Feels Bad When Does Something Wrong	P	-.16	.03	0	-.06	.01	0	-.13	-.03	0	-.16	-.03	0	.68	.79	4
	T	.60	.69	4	.15	.24	0	-.03	-.03	0	-.16	-.03	0			
Concerned About Feelings of Others	P	-.14	.02	0	-.11	.02	0	-.01	.12	0	-.11	-.02	0	.75	.84	4
	T	.59	.67	4	.16	.21	0	.10	.18	0	-.14	-.04	0			
Does Not Show Feelings	P	-.05	.02	0	-.01	.04	0	-.07	.04	0	-.12	-.07	0	.33	.55	3
	T	.14	.32	0	.24	.36	1	-.09	.04	0	-.09	-.31	1			
Keeps the Same Friends	P	-.09	-.02	0	-.06	.03	0	-.03	.08	0	.03	.07	0	.31	.49	3

P=Parent T=Teacher



Table 3

Results of the parent and teacher confirmatory factor analysis

Item	Informant	Min	Max	Item	Informant	Min	Max	Item	Informant	Min	Max
<b>ODD/CD</b>											
Lost Temper	P	.66	.72	<b>Inattention</b>	P	.68	.78	Talked Excessively	P	.56	.63
Argued w/ Adults	T	.87	.90	Careless	T	.81	.86	<b>Relational Regression</b>	T	.80	.85
Defied/Refused	P	.63	.68	Mistakes	P	.71	.80	Excludes to Get Even	P	.76	.82
Deliberately Annoys	T	.90	.93	Difficulty Paying Attention	T	.90	.92	Spreads Rumors	P	.64	.79
Blames Others	P	.62	.70	Not Listening	P	.66	.81	Tries to Get Others to Stop Liking	T	.92	.95
Touchy/Easily Annoyed	T	.90	.91	Failed to Finish	P	.87	.90	Lies About Peers	P	.85	.90
Angry or Resentful	P	.69	.78	Disorganized	T	.73	.84	Ignores	T	.95	.96
Vindictive	P	.92	.94	Avoided Mental Effort	P	.91	.95	Threatens to Stop Being Friend	P	.81	.89
Bullied	T	.67	.78	Lose Things	T	.57	.78	Excludes from Group Activities	T	.95	.97
Cruel to People	T	.91	.94	Easily Distracted	P	.90	.92	<b>Callous-Unemotional</b>	P	.67	.73
Lied to Get Things	P	.94	.85	Forgetful	T	.72	.79	Concerned About Schoolwork	T	.90	.94
Started Physical Fights	T	.73	.85	<b>Hyperactivity-Impulsivity</b>	P	.60	.87	Good at Keeping Promises	P	.77	.81
Serious Violence Testlet	P	.93	.97	Blurted out Answers	T	.87	.90	Feels Bad When Does Something Wrong	P	.93	.94
Feels Bad when does something Wrong	T	.77	.84	Difficulty Waiting Turn	P	.45	.61	Concerned About Feelings of Others	P	.68	.78
Cruelty Testlet	P	.80	.88	Interrupt Activities	T	.82	.87	Does Not Show Feelings	P	.31	.41
Concerned about Feelings of Others	T	.71	.75	Run/Climb	P	.76	.82	Keeps the Same Friends	P	.37	.48
				Difficulty Playing Quietly	T	.93	.96				
				'Driven by Motor'	T	.77	.88				

Note: P=Parent T=Teacher

Table 4

Confirmatory factor analysis fit statistics

	$\chi^2$	df	CFI	TLI	RMSEA
Parent					
Year 1	1084.81***	256	0.897	0.949	0.051
Year 2	1102.77***	224	0.874	0.967	0.057
Year 3	1089.12***	211	0.898	0.969	0.060
Year 4	1013.91***	213	0.902	0.973	0.057
Teacher					
Year 1	711.84***	99	0.932	0.988	0.091
Year 2	996.74***	116	0.946	0.988	0.084
Year 3	1030.75***	115	0.927	0.989	0.087
Year 4	949.37***	111	0.945	0.991	0.085

Note: CFI=comparative fit index; TLI=Tucker-Lewis Index; RMSEA=root mean square error of approximation

Table 5

Correlations between the parent- and teacher-report constructs within time points.

	Parent					Teacher				
	1	2	3	4	5	1	2 <sup>a</sup>	3	4	
Year 1										
1. Inattention	—					—				
2. ODD/CD <sup>a</sup>	.51	—				.61	—			
3. Hyperactivity-Impulsivity	.62	.66	—			.64	.80	—		
4. Relational Aggression	.35	.49	.42	—		.40	.77	.57	—	
5. Callous-Unemotional <sup>b</sup>	.35	.42	.34	.32	—					
Year 2										
1. Inattention	—					—				
2. ODD/CD <sup>a</sup>	.58	—				.62	—			
3. Hyperactivity-Impulsivity	.64	.64	—			.64	.78	—		
4. Relational Aggression	.36	.58	.42	—		.39	.74	.58	—	
5. Callous-Unemotional <sup>b</sup>	.46	.51	.47	.48	—					
Year 3										
1. Inattention	—					—				
2. ODD/CD <sup>a</sup>	.59	—				.61	—			
3. Hyperactivity-Impulsivity	.64	.65	—			.61	.78	—		
4. Relational Aggression	.41	.59	.48	—		.45	.83	.63	—	
5. Callous-Unemotional <sup>b</sup>	.41	.49	.40	.48	—					
Year 4										
1. Inattention	—					—				
2. ODD/CD <sup>a</sup>	.66	—				.70	—			
3. Hyperactivity-Impulsivity	.64	.63	—			.69	.84	—		
4. Relational Aggression	.47	.57	.54	—		.54	.81	.70	—	
5. Callous-Unemotional <sup>b</sup>	.45	.50	.42	.51	—					

<sup>a</sup>For teachers, the ODD/CD factor included two callous-unemotional items.

<sup>b</sup>No callous-unemotional construct exists for teacher data.

Table 6

Intraclass correlations for parent-report and teacher-report constructs across time.

	P/T	1	Year 2	3	Year 1-4 ICC
CD/ODD					
Year 2	P	.781			
	T	.731			
Year 3	P		.805		.878-Parent
	T		.701		.811-Teacher
Year 4	P			.832	
	T			.719	
Inattention					
Year 2	P	.721			
	T	.752			
Year 3	P		.762		.838-Parent
	T		.719		.828-Teacher
Year 4	P			.775	
	T			.730	
Hyperactivity/Impulsivity					
Year 2	P	.779			
	T	.741			
Year 3	P		.780		.862-Parent
	T		.672		.801-Teacher
Year 4	P			.802	
	T			.692	
Relational Aggression					
Year 2	P	.704			
	T	.558			
Year 3	P		.760		.841-Parent
	T		.565		.687-Teacher
Year 4	P			.809	
	T			.566	
Callous-Unemotional Behavior					
Year 2	P	.711			
Year 3	P		.747		.837-Parent
Year 4	P			.755	

Note: P=Parent T=Teacher