Informant, gender and age differences in ratings of adolescent problem behaviour

Jan van der Ende & Frank C. Verhulst

European Child & Adolescent Psychiatry, 2005, 14, 117-126

ABSTRACT

The aim of this study was to determine gender and age differences and agreement among the reports of adolescent problem behaviour by parents, teachers, and adolescents themselves. For 1,122 11-18-year-olds, reports by parents, teachers, and adolescents were obtained. Multivariate regressions were performed to investigate the effect of informant, gender, and age on problem behaviour. Adolescents reported higher levels of problems than parents and teachers for all types of behaviour. Parents reported higher levels of problem behaviour than teachers. Gender differences among informants were dependent on type of problem behaviour. With increasing age, scores of adolescents, parents, and teachers diverged for most types of problems, with larger differences for older adolescents than for younger adolescents. Norms for adolescents need age adjustments for reports by adolescents, parents, and teachers. To obtain a complete view on children's problem behaviours information from multiple informants and differences among them is needed.

INTRODUCTION

The use of multiple informants, for example parents, teachers, and children themselves in child and adolescent psychiatry is acknowledged nowadays (Achenbach, 1991; Rutter & Sroufe, 2000; Verhulst & Van der Ende, 2002), although informants often disagree about problem behaviour (Achenbach, McConaughy, & Howell, 1987; Meyer et al., 2001; Rutter & Sroufe, 2000; Van der Ende, 1999).

Achenbach et al. (1987) reviewed 119 studies that reported associations among at least two informants. They found that among these studies the highest agreement was between the same types of informants. For example, the mean correlation between ratings of mothers and fathers was .59. Although mothers and fathers often observe their children in the same situations, and they may discuss their children's behaviour with each other, this agreement is far from perfect. Differences in ratings by mothers and fathers may be influenced by the amount of time spent with children, interactions with children, and thresholds in reporting problems. The agreement between different types of informants was lower than between the same types of informants. The mean correlation between ratings of parents and teachers was .27, between ratings of parents and their children .25, and between ratings of children and their teachers .20. Agreement is higher for informants that observe and interact with children in the same situations, e.g., mothers and fathers, than for informants in different situations, e.g., parents and teachers. Correlations were higher for 6- to 11-year-olds than for 12- to 16-year-olds, and correlations were higher for externalising than for internalising problems. Few studies are available that report on other variables that moderate agreement.

The low correlation among the ratings of different types of informants suggests that each type of informant contributes its own set of information about a child's behaviour, and that information from one type of informant cannot replace the information from another type of informant. Any single assessment method provides only a partial description of a child's behaviour (Achenbach et al., 1987; Meyer et al., 2001). Recent studies confirmed the low agreement among informants of ratings of problem behaviour (Meyer et al., 2001), and also showed that informants are not interchangeable (Rubio-Stipec, Fitzmaurice, Murphy, & Walker, 2003).

Only a few studies are available that showed differences in level of ratings among informants. Most studies that compared ratings of informants used different methods to assess behaviour problems, e.g., rating scales and interviews. Unequal assessment methods may hinder direct comparison of ratings, because they may differ in assessment procedure, problem content, item format, or scaling.

The past decades assessment instruments emerged that have similar versions for informants, mostly parents, teachers, and children themselves (Verhulst & Van der Ende, 2002). Studies that reported on differences among informants obtained with similar

assessment instruments showed that adolescents themselves reported higher levels of problem behaviour than their parents and teachers, and that parents reported higher levels than teachers (Sawyer, Baghurst, & Mathias, 1992; Seiffge-Krenke & Kollmar, 1998; Stanger & Lewis, 1993; Verhulst & Van der Ende, 1992). Most studies compared only two informants. When studies compared more than two informants only pairwise comparisons were made. Studies directly comparing informants in one analysis are scarce.

Norms based on large samples allow clinicians to compare a patient with a relevant group of subjects (Meyer et al., 2001). A clinician can decide how much an individual patient deviates from a reference group. For assessment instruments that have versions for multiple informants, separate norms for each informant are reported. It is also important to have information on differences among informants as from large reference samples (Meyer et al., 2001; Rutter & Sroufe, 2000; Van der Ende, 1999). A difference that deviates from what is generally found, may guide clinicians in choosing next steps, e.g. interview with parent, type of treatment. Achenbach (1991) provided information about agreement between two informants in large reference samples. However, norms on differences among informants are not available yet, and are therefore needed.

This study uses a sizable (n=1,122) sample from the general population with ratings of parents, teachers, and adolescents, on comparable, standardized assessment instruments for assessing a broad range of child and adolescent behavioural and emotional problems. In this study we aim to determine: (a) systematic differences in the level of ratings among informants; (b) dependencies of differences among informants on gender, age, and SES; (c) agreement of ratings among informants; (d) dependencies of agreement among informants on gender, age, and SES.

METHOD

Sample and procedure

The target population consisted of all 4- to 18-year-olds who were of Dutch nationality and were living in the Netherlands on January 1, 1993. The population was stratified according to four country regions and four degrees of urbanization. A two-stage sampling procedure was followed, with a random selection of municipalities (n=89) in the first, followed by a random selection of individuals equally divided across gender and age in the second stage. Trained lay interviewers with previous survey experience visited or phoned the parent to make an appointment, preferably with the mother. Of the 2,709 subjects for whom informants were reached, parent interviews were completed for 2,227 (82.2%). A more detailed description of the sample selection and methodology is reported elsewhere (Verhulst, Van der Ende, Ferdinand, & Kasius, 1997).

In this study only subjects of 11 years and older were included. For 1,122 subjects, reports from 1122 parents, reports from 818 teachers, and 1,114 self-reports were obtained. Of the 1,122 subjects 557 were boys. Their mean age was 14.4 years. There were 565 girls, and their mean age was 14.5 years.

Of the 1,122 parent reports, 1,056 (94.9%) reports were provided by mothers, 52 (4.6%) by fathers, and 5 (.4%) by others. For the 1,122 subjects in this study 304 (37.2%) teacher reports could not be obtained. The reasons for not obtaining teacher reports were: Parents refused permission to contact teachers (n=69, 23.7%); adolescents were not in school anymore (n=35, 11.5%); teachers refused or did not respond (n=200, 65.8%). Subjects for whom no teacher reports were available did not differ significantly in gender (χ^2=.07, df=1, p>.05) and parent reported problem behaviour (t=1.83, df=1,120, p>.05) from subjects for whom teachers reports were available. However, subjects with missing teachers' reports were about 1 year older (t=6.80, df=1,120, p<.05) than subjects for whom teacher reports were available.

Instruments

The Child Behavior Checklist (CBCL), Teachers' Report Form (TRF), and Youth Self-Report (YSR) were used to obtain standardized reports of adolescents' problem behaviours from parents, teachers, and adolescents themselves respectively (Achenbach, 1991). The good reliability and validity of the original American instruments were confirmed for the Dutch translations (Verhulst et al., 1997).

In this study only the items about problem behaviour were used. These items can be scored on eight scales. The content of the scales is the same for all three informants, although the scales may vary in number of items across informants. The scales are labelled Withdrawn, Somatic Complaints, Anxious/Depressed, Social Problems, Thought Problems, Attention Problems, Delinquent Behavior, and Aggressive Behavior. The items can be further aggregated into the scales Internalizing, Externalizing, and Total Problems.

Socioeconomic status (SES) was indicated by a 6-level scale of occupation (6=high) (Westerlaak, Kropman, & Collaris, 1975). The scale was dichotomised into low status (scores 1–3) and high status (scores 4–6).

Statistical Analysis

Scales of the CBCL, TRF, and YSR differ in item composition and item number. To compare the scales we included only the 89 items that were common to all three instruments. In this way the same set of behaviours was rated by the informants. The total raw scores of the composed scales were included in the analyses.

Multivariate regressions were used to analyse the effect of gender, age and SES on the differences among informants of problem behaviour (Goldwasser & Fitzmaurice, 2001). A vector consisting of the responses from the three informants was used as the multi-

variate dependent variable. This method adjusts for the correlation between the ratings of informants and also handles missing data by including all available reports of 1,122 subjects comprising 1,122 parent reports, 818 teacher reports, and 1,114 self-reports. Tests of interactions between the variable that indicates the type of informant and gender, age, and SES were used to provide evidence for the effect of gender, age, and SES on informant differences. Graphs of the difference among informants were constructed to facilitate the interpretation of the results. A quadratic age factor, denoted as age^2, was also tested, because previous research provided support for these effects (Bongers, Koot, van der Ende, & Verhulst, 2003). We proceeded analyses by first testing for interactions and subsequently testing more simple models. Multivariate regressions were performed with the SAS PROC MIXED procedure, version 8.2.

Correlations were used to analyse agreement among the informants. Separate analyses were performed for boys versus girls and younger (11-14 years) versus older (15-18 years) adolescents.

RESULTS

Multivariate regressions

Table 1, and figures 1 and 2 show the final models as result of the multivariate regressions. SES is not displayed in the graphs, because SES was a main effect in all analysis, but no interactions between SES and informant, gender, and age were significant. Subjects with low SES scored higher on all scales than subjects with high SES. We focussed on informant differences and used higher order interactions as a starting point for interpretation, because these interactions incorporated lower order interactions and main effects.

Only the interaction of informant by gender was significant for Withdrawn. This indicated that the difference between boys and girls was larger for adolescents themselves than for adult informants. Also, scores of adolescents themselves were higher for girls than for boys, whereas scores of adult informants were higher for boys than for girls.

For Somatic Complaints the interaction of informant by gender was significant. This indicated that the difference between boys and girls was larger for adolescents themselves than for adult informants. The difference between boys and girls was larger for parents than for teachers.

For Anxious/Depressed, two interactions involving informant were significant: Informant by gender, and informant by age. These interactions indicated that the differences between boys and girls were larger for adolescents themselves than for adult informants. Differences between informants changed across age. Scores of adolescents themselves increased across age, especially for girls, whereas scores of adult informants decreased across age, especially for boys.

Table 1. Significant effects of informant, gender, age, and SES in final models from multivariate regressions.

	Scales										
	Withdrawn	**Somatic Complaints**	**Anxious/ Depressed**	**Social Problems**	**Thought Problems**	**Attention Problems**	**Delinquent Behavior**	**Aggressive Behavior**	**Internalizing**	**Externalizing**	**Total Problems**
Effects	F	F	F	F	F	F	F	F	F	F	F
Informant (2)	72.8	399.7	15.3	141.5	22.4	9.2	77.5	16.2	45.2	24.6	47.5
Gender (1)	*.1*	48.0	.3	9.3	.6	8.1	23.4	20.5	22.9	15.5	4.5
Age (1)	12.4		*.3*	8.6	*.7*	8.4	21.8	5.0	7.2	10.3	8.9
SES (1)	9.4	5.5	5.3	9.2	5.5	9.4	5.9	14.1	10.0	12.9	18.6
A^2 (1)	4.8					5.9	7.5	6.5	4.6	8.2	7.3
I X G (2)	3.0	10.1	5.4	3.7	5.3	13.4		8.0	7.2	6.9	7.2
I X A (2)			11.3		6.9	23.5	33.8	9.9	6.2	11.8	8.0
G X A (1)			5.9		*1.7*	*.4*		4.3		*1.3*	*2.4*
I X A^2 (2)								4.1		4.1	3.7
I X G X A (2)					3.1	4.1		5.4		5.4	4.4
DF Error	1118	1120	1118	1119	1118	1117	1118	1118	1118	1117	1117

All F-values are significant at the $p < .05$ level, except those in italics; these effects are included because higher order interactions were significant. Degrees of freedom for the effects are between parentheses. Full descriptions of the interactions are: I X G: Informant by Gender; I X A: Informant by Age; G X A: Gender by Age; I X A: Informant by Age^2; I X G X A: Informant by Gender by Age.

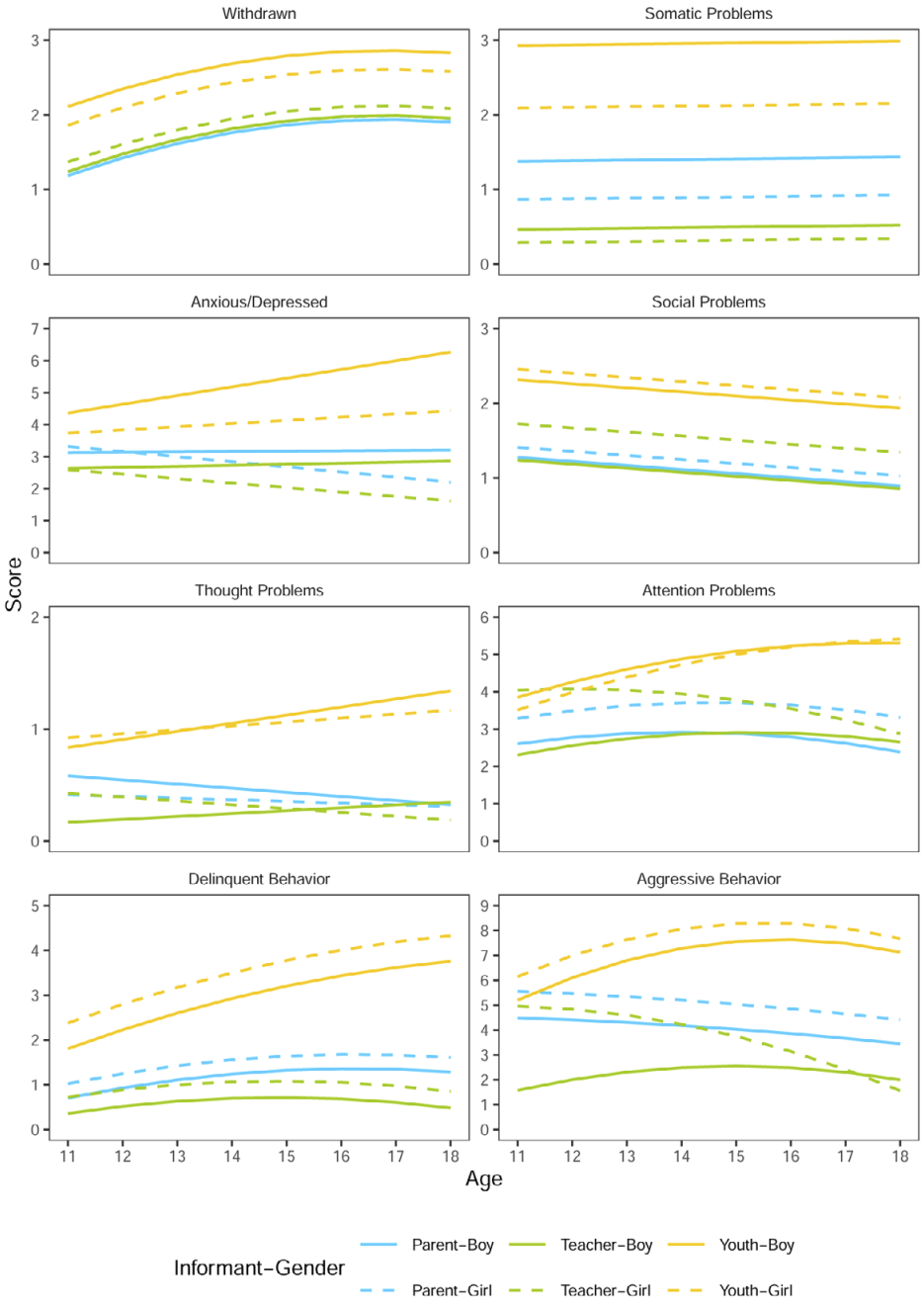

Figure 1. Graphical display of differences among informants. Y-axis displays means of the small-band scale scores of the CBCL, TRF, and YSR separate for boys and girls.

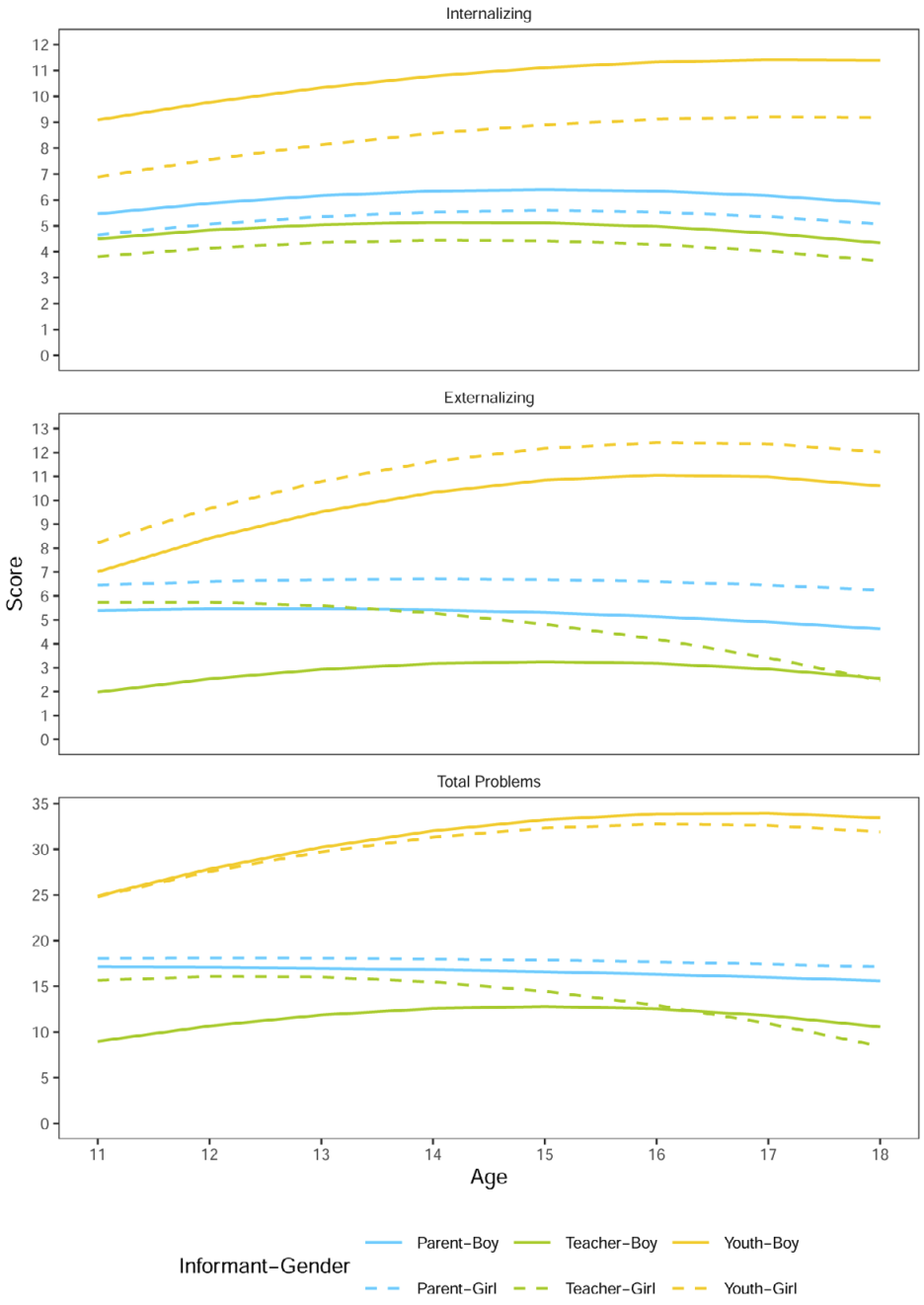

Figure 2. Graphical display of differences among informants. Y-axis displays means of the broad-band scale scores of the CBCL, TRF, and YSR separate for boys and girls.

Table 2. Pearson correlations between CBCL, TRF, and YSR scales based on items common to all three instruments

	Parents with Teachers				Parents with Adolescents				Teachers with Adolescents				
	Boys		Girls		Boys		Girls		Boys		Girls		
	11-14	15-18	11-14	15-18	11-14	15-18	11-14	15-18	11-14	15-18	11-14	15-18	
	(n=228)	(n=180)	(n=211)	(n=199)	(n=280)	(n=273)	(n=270)	(n=291)	(n=226)	(n=179)	(n=210)	(n=199)	Mean
Withdrawn	.41	.24	.33	.42	.43	.25	.34	.35	.23	.21	.19	.32	.31
Somatic Complaints	.14	*.10*	.28	.20	.26	.25	.28	.38	.15	*.12*	.20	*.06*	.21
Anxious/Depressed	.33	*.13*	.38	.25	.38	.37	.26	.35	.16	.16	.25	.16	.27
Social Problems	.42	.25	.22	.43	.28	.36	.27	.30	.24	.25	.30	.33	.30
Thought Problems	.15	.16	.23	.13	.25	*.09*	.12	.19	*.10*	*-.06*	.15	.19	.14
Attention Problems	.41	.29	.38	.39	.32	.40	.39	.43	.27	.26	.43	.29	.35
Delinquent Behavior	.34	.55	.37	.47	.18	.45	.33	.41	*.11*	.39	.28	.42	.36
Aggressive Behavior	.36	.35	.37	.43	.26	.38	.42	.43	.15	.26	.49	.35	.35
Internalizing	.35	.19	.41	.34	.44	.38	.28	.35	.23	.19	.23	.19	.30
Externalizing	.38	.46	.41	.49	.25	.44	.43	.46	*.12*	.32	.49	.43	.39
Total Problems	.41	.30	.39	.45	.35	.37	.31	.39	.22	.21	.36	.29	.34
Mean	.31	.25	.31	.33	.28	.31	.28	.33	.16	.19	.28	.25	.28

Correlations are significant at $p < .05$, except for the correlations in italics (9 out of 132).

The only interaction for Social Problems, informant by gender, indicated that the difference between boys and girls was larger for teachers than for adolescents themselves and parents.

The results for Thought Problems revealed an interaction for informant by gender by age. The scores of adolescents themselves increased across age. Boys had higher scores than girls, but scores of girls increased more rapidly across age. Scores of parents decreased across age for both boys and girls. Scores of teachers decreased for boys and increased for girls.

For Attention Problems there was an interaction for informant by gender by age. Most striking were the small differences among the informants at age 11 and the large differences at age 18.

For Delinquent Behavior the interaction of informant by age indicates increasing informant differences across age.

For Aggressive Behavior two interactions, informant by gender by age, and informant by age^2 were significant. The scores of teachers for boys dropped strikingly across age. Differences between adolescents' reports and parents' reports increased with age.

The two significant interactions for Internalizing, informant by gender and informant by age, indicated that differences between boys and girls were larger for self-reports than for adult informants' reports and that differences among informants increased with age.

For Externalizing two interactions, informant by gender by age and informant by age^2 were significant. These interactions indicated that scores of adolescents themselves increased and scores of adult informants decreased with age, especially scores of teachers for boys.

For Total Problems two interactions, informant by gender by age and informant by age^2 were significant, indicating increasingly higher scores of adolescents themselves and lower scores of adult informants with age, especially scores of teachers for boys.

The results for Aggressive Behavior, Externalizing, and Total problems revealed a drop across age in scores of teachers for boys that could be the result of the 304 missing teacher reports. Therefore, we tested all final models on the 818 subjects with complete data, i.e. reports of all three informants were available. Although a few higher order interactions were not significant anymore, due to less power, graphs constructed from the results of the analyses of the complete data were similar to the reported graphs and revealed identical interpretations. This was also true for graphs that showed a drop in scores of teachers for boys.

Correlations

Correlations between pairs of informants are shown in table 2. Correlations are given for each scale, gender, and age groups (11–14 years versus 15–18 years). The last row in table 2 shows the mean correlations for gender and age groups. The last column shows the mean correlation for each scale.

Most correlations were significant, but many of them were low. The mean correlations for the two age groups within boys and the pair teacher and adolescents were .16 and .19. These correlations were the lowest among the mean correlations for gender and age groups within pairs of informants. The other gender and age groups did not show much variation in the mean correlations. The lowest mean correlation for scales was .14 for Thought Problems. The highest correlations were for the scales Attention Problems, Delinquent Behavior, and Aggressive Behavior.

DISCUSSION

This study showed that the direct comparison of ratings of problem behaviour by parents, teachers, and adolescents revealed differences among informants that depend on gender, age, and type of problem. Adolescents generally scored themselves higher than parents and teachers. Parents generally scored adolescents higher than teachers, although these differences were not as striking as the difference in scores by adolescents themselves versus their parents and teachers. Previous studies also reported differences among informants (Sawyer et al., 1992; Seiffge-Krenke & Kollmar, 1998; Stanger & Lewis, 1993; Verhulst & Van der Ende, 1992), but samples used in these studies were generally small or only two informants were directly compared. Our results, i.e., differences among informants, with larger differences with increasing age, and uneven differences between boys and girls, have implications for the interpretation of ratings of different informants which will be discussed below.

We also found an effect for socioeconomic status (SES), but this effect did not depend on informant, gender, or age. These findings support the evidence that low SES in families increases the risk for psychopathology in children (Keiley, Bates, Dodge, & Pettit, 2000; Ritsher, Warner, Johnson, & Dohrenwend, 2001). We found only low to moderate correlations of ratings among informants. These correlations commensurate with or are only little higher than correlations reported earlier (Achenbach et al., 1987; Meyer et al., 2001). Using scales that included only items that were common to all three informants did not lead to a substantial increase in agreement. Disagreement among informants has to be taken as a fact and differences among informants should be taken into account when evaluating children for clinical purposes or in research.

Gender and age differences among informants

For all scales gender or age or both were involved in interactions with informant. Thus, differences between boys and girls, and differences across age in types of problems are dependent on the informant that provides information on problems of adolescents. Gender and age differences are well recognized for several types of problems (Rutter &

Sroufe, 2000), but understanding the development of gender differences in childhood and adolescence is hampered by the paucity of available descriptive data on gender and age differences and is complicated by variations in sampling, statistical analysis, and informants (Lahey et al., 2000).

Internalising problems

Previous research indicated that internalising problems are more prevalent in girls than in boys and that for depressive symptoms the difference between boys and girls is getting larger with increasing age (Angold, Costello, & Worthman, 1998; Nolen-Hoeksema & Girgus, 1994). Our results revealed interactions of informant by gender and informant by age for Anxious/Depressed, and Internalizing, and interactions of informant by age only for Withdrawn and Somatic Complaints.

Differences in depressive symptoms between boys and girls emerge in adolescence (Angold et al., 1998; Nolen-Hoeksema & Girgus, 1994). However, most studies reporting gender difference in depressive symptoms relied on self-reports only (Nolen-Hoeksema & Girgus, 1994). Our results confirmed this increasingly larger gender difference of depressive symptoms for self-reports of adolescents. Scores of the Anxious/Depressed scale increased for girls at a higher rate than for boys with increasing age. However, both adult informants revealed a different pattern. Parents and teachers also showed an increasingly larger gender difference, but now scores of girls stay at the same level and scores of boys decrease with increasing age. Withdrawn and Somatic Complaints lacked this informant by age interaction. The items on the Anxious/Depressed scale pertain to thoughts, feelings, and moods, for example "feels unloved" and "feels worthless" that are not easily observable for adult informants. As adolescents grow older adult informants may be less aware of these problems, because they have less contact with adolescents or adolescents talk less often about their problems with their parents and teachers. Withdrawn and Somatic Complaints pertain to problems that are more easily observable than for Anxious/Depressed. For example "refuses to talk" and "underactive" are problems on Withdrawn and "headaches" and "overtired" are problems on Somatic Complaints. Informant differences are evident for these scales. Self-reports show higher scores than adult reports, but with increasing age this difference remains.

Externalising problems

Previous research indicated that externalising problems are more prevalent in boys than in girls (Lahey et al., 2000; Zahn-Waxler, 1993; Zoccolillo, 1993). Aggressive conduct problems are more prevalent in younger adolescents and non-aggressive conduct problems are more prevalent in older adolescents (Lahey et al., 2000). In our study we found for both Delinquent Behavior and Aggressive Behavior an interaction of informant by age. For Aggressive Behavior we found an interaction between informant and gender, but

Delinquent behavior was the only scale for which we did not find a significant interaction of informant by gender.

Delinquent Behavior, comprised of items as "lies" and "truancy", may be denoted as non-aggressive conduct problems and Aggressive Behavior, comprising of items as "fights" and "attacks people", may be denoted as aggressive conduct problems. For self-reports earlier studies on non-aggressive conduct problems were confirmed by our results. Scores of Delinquent Behavior were higher for boys than for girls and scores were increasingly higher with increasing age. However, parents and teachers did not show an increase as strong as for adolescents themselves. Several explanations pertain to this differences between self-reports and adult informant reports. The increase is real, but parents and teachers are less aware of these behaviours, because adolescents conceal them to avoid negative consequences. In face to face interviews or mailing surveys, adolescents themselves are willing to report these behaviours. Also, adult informants may be reluctant to admit these socially undesirable behaviours, but since teachers report less problems than parents and teachers are probably less bothered by social undesirability this reason does not deserve much support. For aggressive conduct problems previous studies that used parent reports were confirmed by our results, i.e., scores on Aggressive Behavior were higher for boys than for girls and scores decreased with increasing age (Lahey et al., 2000). However, scores of adolescents themselves increased from age 11 and reached a peak about age 15 and decreased from age 16, but this latter age interval may have been too short to decide if this decreasing trend will continue into young adulthood. Most puzzling were the results for teacher reports. Boys and girls had different scores at age 11 and scores were equal around age 16 through age 18. This result supports the model of Silverthorn and Frick (1999) who hypothesized that girls engage in conduct problems at a later age than boys. But since only reports by teachers showed converging developmental paths for boys and girls in adolescence, and converging paths are more a result of decreasing scores for boys than increasing scores for girls, this conclusion must be taken with caution and more research is needed to explore this phenomenon.

Other problems

The Social Problems scale comprises problems like "gets teased" and "not liked by peers". Older adolescents are scored lower than younger adolescents by all three informants. But differences between boys and girls were larger for reports by teachers than for reports by parents and self-reports. Teachers can observe social behaviour among their pupils in the classroom. They can observe many interactions of boys and girls with each other. They have large samples of social behaviour available upon which they can base their ratings of social problems. Therefore, teachers can better discriminate between boys and girls, hence the difference in ratings by teachers, because they have more information available. Parents can observe social behaviour among their own children or when their

children are with their peers. Adolescents themselves will probably report on their social problems they encounter when they are with their family members or with their close friends. Therefore, parents and adolescents will often report on a smaller sample of social behaviours than teachers do.

The Thought Problems scale comprises a mixture of obsessions and compulsions like "repeats acts" and psychotic signs like "sees things". These problems are often not well observable by others. Adult informants depend on the information about these problems adolescents reveal to them. Our results showed that older adolescents probably reveal less thought problems to their parents and teachers than younger adolescents.

The Attention Problems scale comprises of problems that pertain to inattention, like "can't concentrate" and hyperactive behaviour, like "can't sit still". Ratings by teachers showed higher scores for boys than for girls for younger adolescents, but for older adolescents the difference between boys and girls was smaller. This is in agreement with what is known of the epidemiology of attention problems and hyperactive behaviour (Scahill & Schwab-Stone, 2000). However, parent reports showed constant levels of Attention Problems with higher scores for boys than for girls. Self-reports showed an increase in scores of the Attention Problems scale from age 11 through age 18, and probably into young adulthood. The results for parent reports and self-reports lend support to the theory that inattention and hyperactive behaviour may continue into adulthood (Faraone et al., 2000).

Limitations

We constructed scales comprising items common to the CBCL, TRF, and YSR to ensure that all three informants rated the same problems. Differences among informants may be different for the adapted versus original scales. However, using scales with items common to all informants, i.e., each informant rated the same type of problems is an important criterion in assessing sources of bias (Richters, 1992). Two other criteria that are important in assessing sources of bias are temporal assessment of behaviour and situational concordance of assessment. The criterion of temporal assessment is fulfilled in our study because the CBCL, TRF, and YSR were all collected in a short time interval. The criterion of situational concordance is not fulfilled because informants reported on behaviour in different settings. Parents reported on problems at home, teachers reported on problems in school, and adolescents themselves reported on problems at home, in school, and elsewhere.

The generalisability of this study is limited to adolescents in the general population. Results may be different for younger children or for children in clinical samples (Sawyer et al., 1992). Furthermore, since our sample is from the Dutch population, our results may not be generalisable to samples from other countries. However, recent cross-cultural research on adolescents' parent-reported and self-reported problems from twelve (Crijnen,

Achenbach, & Verhulst, 1997) and seven (Verhulst et al., 2003) countries respectively, provides ample support for our results, although ratings of parents and adolescents themselves were not directly compared and ratings of teachers were not included.

Implications

Our results suggest a number of practical implications. First, given the differences among informants and dependencies of these differences on gender and age, our results support recent claims that using a single informant to obtain information will lead to an incomplete or biased understanding of a patient (Meyer et al., 2001). By excluding informants particular aspects of disorders may be missed, so that disorders are not well represented by the reports of included informants (Cole, Truglio, & Peeke, 1997).

Second, norms for adolescents need adjustments for age for self-reports, parents' and teachers' reports. We made use of the CBCL, TRF, which have separated norms for gender and two age groups, 4-11 years and 12-18 years, the YSR, which has only separated norms for gender for 11-18-year-olds (Achenbach, 1991). However, as our results have demonstrated, differences among informants are larger for older adolescents than for younger adolescents. When norm-based scores by informants are compared, differences between ratings of, for example, parents and adolescents themselves will be underestimated for older and overestimated for younger adolescents. Incorporating age adjustments in the norms for adolescents will remove this bias.

Third, reporting of only total scales of rating scales is not always sufficient. As can be seen in figure 2, the graph of Internalizing, is similar to the graph of Withdrawn in figure 1. The graph of Externalizing is similar to the graph of Aggressive Behavior, and that of Total Problems is similar to Externalizing. It appears that total scales reflect mostly one scale that is part of the scales that comprise the total scales. For example, Aggressive Behavior is a large scale with more items and more frequent endorsed items than in Delinquent Behavior, and will therefore put more weight to the Externalizing scale. The distinct results for the lower order scales, e.g., Withdrawn, Anxious/Depressed, Delinquent Behavior, and Aggressive Behavior, justifies reporting of these scales and reporting only higher order scales or total scales, e.g., Internalizing, Externalizing, and Total Problems, will inevitably result in loss of information.

Fourth, because gender and age differences in levels of problems depend on type of informant, averaging scores from multiple informants as is done in some studies (Van der Ende, 1999) may not be optimal. Data analytic methods that account for differences among informants and in which effects of risk factors on informant differences can be tested, as in the current study, lead to more readily interpretable results.

REFERENCES

Achenbach, T. M. (1991). *Integrative guide for the CBCL/4-18, YSR, and TRF profiles*. Burlington, VT: University of Vermont, Department of Psychiatry.

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*(2), 213-232.

Angold, A., Costello, E. J., & Worthman, C. M. (1998). Puberty and depression: the roles of age, pubertal status and pubertal timing. *Psychological Medicine, 28*(1), 51-61.

Bongers, I. L., Koot, H. M., van der Ende, J., & Verhulst, F. C. (2003). The normative development of child and adolescent problem behavior. *Journal of Abnormal Psychology, 112*(2), 179-192.

Cole, D. A., Truglio, R., & Peeke, L. (1997). Relation between symptoms of anxiety and depression in children: A multitrait-multimethod-multigroup assessment. *Journal of Consulting & Clinical Psychology, 65*(1), 110-119.

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*(9), 1269-1277.

Faraone, S. V., Biederman, J., Spencer, T., Wilens, T., Seidman, L. J., Mick, E., & Doyle, A. E. (2000). Attention-deficit/hyperactivity disorder in adults: an overview. *Biological Psychiatry, 48*(1), 9-20.

Goldwasser, M. A., & Fitzmaurice, G. M. (2001). Multivariate linear regression analysis of childhood psychopathology using multiple informant data. *International Journal of Methods in Psychiatric Research, 10*, 1-10.

Keiley, M. K., Bates, J. E., Dodge, K. A., & Pettit, G. S. (2000). A cross-domain growth analysis: externalizing and internalizing behaviors during 8 years of childhood. *Journal of Abnormal Child Psychology, 28*(2), 161-179.

Lahey, B. B., Schwab-Stone, M., Goodman, S. H., Waldman, I. D., Canino, G., Rathouz, P. J., . . . Jensen, P. S. (2000). Age and gender differences in oppositional behavior and conduct problems: a cross-sectional household study of middle childhood and adolescence. *Journal of Abnormal Psychology, 109*(3), 488-503.

Meyer, G. J., Finn, S. E., Eyde, L. D., Kay, G. G., Moreland, K. L., Dies, R. R., . . . Reed, G. M. (2001). Psychological testing and psychological assessment. A review of evidence and issues. *American Psychologist, 56*(2), 128-165.

Nolen-Hoeksema, S., & Girgus, J. S. (1994). The emergence of gender differences in depression during adolescence. *Psychological Bulletin, 115*(3), 424-443.

Richters, J. E. (1992). Depressed mothers as informants about their children: a critical review of the evidence for distortion. *Psychological Bulletin, 112*(3), 485-499.

Ritsher, J. E., Warner, V., Johnson, J. G., & Dohrenwend, B. P. (2001). Inter-generational longitudinal study of social class and depression: a test of social causation and social selection models. *British Journal of Psychiatry Supplement, 40*, s84-90.

Rubio-Stipec, M., Fitzmaurice, G., Murphy, J., & Walker, A. (2003). The use of multiple informants in identifying the risk factors of depressive and disruptive disorders--are they interchangeable? *Social Psychiatry & Psychiatric Epidemiology, 38*(2), 51-58.

Rutter, M., & Sroufe, L. A. (2000). Developmental psychopathology: concepts and challenges. *Development and Psychopathology, 12*(3), 265-296.

Sawyer, M. G., Baghurst, P., & Mathias, J. (1992). Differences between informants' reports describing emotional and behavioural problems in community and clinic-referred children: a research note. *Journal of Child Psychology & Psychiatry, 33*(2), 441-449.

Scahill, L., & Schwab-Stone, M. (2000). Epidemiology of ADHD in school-age children. *Child and Adolescent Psychiatric Clinics of North America, 9*(3), 541-555.

Seiffge-Krenke, I., & Kollmar, F. (1998). Discrepancies between mothers' and fathers' perceptions of sons' and daughters' problem behaviour: a longitudinal analysis of parent- adolescent agreement on internalising and externalising problem behaviour. *Journal of Child Psychology & Psychiatry, 39*(5), 687-697.

Silverthorn, P., & Frick, P. J. (1999). Developmental pathways to antisocial behavior: the delayed-onset pathway in girls. *Development and Psychopathology, 11*(1), 101-126.

Stanger, C., & Lewis, M. (1993). Agreement among parents, teachers, and children on internalizing and externalizing behavior problems. *Journal of Clinical Child Psychology, 22*(1), 107-115.

Van der Ende, J. (1999). Multiple informants: multiple views. In H. M. Koot, A. A. M. Crijnen, & R. F. Ferdinand (Eds.), *Child psychiatric epidemiology* (pp. 39-52). Assen: Van Gorcum.

Verhulst, F. C., Achenbach, T. M., van der Ende, J., Erol, N., Lambert, M. C., Leung, P. W., . . . Zubrick, S. R. (2003). Comparisons of problems reported by youths from seven countries. *American Journal of Psychiatry, 160*(8), 1479-1485.

Verhulst, F. C., & Van der Ende, J. (1992). Agreement between parents' reports and adolescents' self-reports of problem behavior. *Journal of Child Psychology & Psychiatry, 33*(6), 1011-1023.

Verhulst, F. C., & Van der Ende, J. (2002). Rating scales. In M. Rutter & E. Taylor (Eds.), *Child and Adolescent Psychiatry* (4th ed., pp. 70-86). London: Blackwell Publishing.

Verhulst, F. C., Van der Ende, J., Ferdinand, R. F., & Kasius, M. C. (1997). The prevalence of DSM-III-R diagnoses in a national sample of Dutch adolescents. *Archives of General Psychiatry, 54*(4), 329-336.

Westerlaak, J. M., Kropman, J. A., & Collaris, J. W. M. (1975). *Beroepenklapper*. Nijmegen: Instituut voor Toegepaste Sociologie.

Zahn-Waxler, C. (1993). Warriors and worriers: Gender and psychopathology. *Development and Psychopathology, 5*(1-2), 79-89.

Zoccolillo, M. (1993). Gender and the development of conduct disorder. *Development and Psychopathology, 5*(1-2), 65-78.