

## **Differential Predictive Value of Parents' and Teachers' Reports of Children's Problem Behaviors: A Longitudinal Study**

**Frank C. Verhulst,<sup>1,2</sup> Hans M. Koot,<sup>1</sup> and Jan Van der Ende<sup>1</sup>**

*This study investigated the prediction of signs of disturbance in 946 children originally aged 4 to 11 years from the general population across a 6-year period. Parents' and teachers' ratings obtained via the Child Behavior Checklist (CBCL) and Teacher's Report Form (TRF) were tested as predictors of (a) academic problems, (b) school behavior problems, (c) receipt of mental health services, (d) child's need for professional help, (e) suicidal behavior, and (f) police contacts. Total problem scores in the deviant range on the CBCL or TRF were significantly associated with poor outcomes 6 years later. The combination of deviant scores on both the CBCL and TRF was a powerful predictor of poor outcomes with 56% of the girls, and 36% of the boys with total problem scores in the deviant range on both instruments maladjusted 6 years later. The CBCL syndromes Attention Problems and Delinquent Behavior, and the TRF syndromes Delinquent Behavior, Somatic Complaints, and Social Problems significantly predicted poor outcomes. Teachers' reports predicted poor outcomes equally well or even somewhat better than parents' reports. It is important to include teacher information in the diagnostic assessment of children.*

Assessment of psychopathology in children should ideally be based on information from multiple sources. Correlations between reports from different informants tend to be low, averaging .28 between informants who

Manuscript received in final form February 8, 1994.

We wish to thank Dr. Thomas Achenbach for his helpful advice. This research was supported by grants from the Sophia Foundation for Medical Research, and the Health Research Promotion Program.

<sup>1</sup>Department of Child and Adolescent Psychiatry and Sophia Children's Hospital, Erasmus University Rotterdam, Rotterdam, The Netherlands.

<sup>2</sup>Address all correspondence, including reprint requests, to Frank C. Verhulst, M.D., Sophia Children's Hospital/Erasmus University Rotterdam, Department of Child and Adolescent Psychiatry, Dr. Molewaterplein 60, 3015 GJ Rotterdam, The Netherlands.

see children in different contexts such as parents and teachers (Achenbach, McConaughy, & Howell, 1987). These low correlations reflect the situational variability of children's behavior as well as differences in informant characteristics and differences in informant-child interactions. Because each informant may provide valid, though different, information, reports from different sources are needed. For school-aged children, it is usually the parents and teachers who provide relevant information, supplemented with information from the clinical assessment of the child.

Although the need for multiple informants has long been stressed (Achenbach, 1982, 1985), little empirical knowledge is available on how to handle information from different sources. Different informants may provide either discrepant or corresponding information. This has been treated by some authors as a nuisance forcing diagnosticians to decide on who is the best informant rather than viewing variability in the level of agreement as a possible source of valuable information.

Several studies have tried to determine which informant is best for certain disorders. Reich and Earls (1987) reported rather arbitrary decision rules for how to weigh information from different sources to obtain certain diagnoses. Loeber, Green, and Lahey (1990) reported on clinicians' preference for particular informants regarding different areas of psychopathology, but did not test the validity of these opinions.

In a sample of boys referred for disruptive behaviors, Loeber, Green, Lahey, and Stouthamer-Loeber (1990) computed the "conditional agreement" for parents', teachers', and children's reports to determine the adequacy of reports derived from these informants and to select the optimal informant for a certain disorder. For instance, if many children failed to report a particular problem behavior that their parents reported about them, the conditional probability of children reporting their problem behavior is low, given that their parents reported the behavior. In such cases, the authors concluded that the information from the parent is necessary for the assessment of that particular problem behavior. However, information provided by one informant, but not by the other, need not necessarily be valid, as was implicitly assumed by the authors.

Using standardized parent and child interviews as a systematic way of eliciting information on problem behavior in an epidemiological sample of children aged 9 to 16 years, Bird, Gould, and Staghezza (1992) selected the informant whose interview responses were the most predictive for a clinical diagnosis. The same child psychiatrist who coded the responses for both the parent and child interviews also provided the clinical diagnoses, which were based on interview information and a review of teacher reports. The results may therefore reflect the clinician's preference for a certain

informant in making the diagnosis. There was no external criterion to test the validity of the clinician's diagnosis.

One way to test the validity of information from different sources is to evaluate the power to predict maladjustment longitudinally. A number of epidemiological studies have shown that children's problem behaviors assessed in a standardized way are significantly correlated across time (McConaughy, Stanger, & Achenbach, 1992; Verhulst, Koot, & Berden, 1990; Verhulst & Van der Ende, 1991a, 1992a, 1992b). Correlations were higher for ratings obtained from the same type of informant at each assessment, versus ratings obtained from different types of informants. For instance, the 3-year stability for parent ratings reported by McConaughy et al. (1992) was .58 for total problem scores in a representative U.S. national sample, whereas the correlation for total problem scores obtained from parent reports at Time 1 and self-reports at Time 2 was .22, and for total problem scores from parent reports at Time 1 and teacher reports at Time 2, it was .25.

Significant stability of ratings obtained from the same individual (e.g., the same parent on different occasions) reflects the stability of the child's behavior in a certain environment (e.g., the home), as well as stability in individual rater characteristics and rater-child interaction characteristics. If ratings on different occasions are obtained from the same type of informant, but from different individuals at different assessments (e.g., different teachers on different occasions), the stability is lower because stability in individual rater characteristics and specific rater-child interaction characteristics are eliminated. These effects may have influenced the 4-year stability of Child Behavior Checklist parent ratings ( $r = .66$  for total problem scores; Verhulst et al., 1990), which was greater than that for teacher ratings using the Teacher's Report Form across the same interval ( $r = .37$  for total problem scores; Verhulst & Van der Ende, 1991a).

These epidemiological studies were informative with respect to the typical stabilities of children's problem behaviors assessed within or across different types of informants over time. However, these studies did not allow the comparison of different types of informants with respect to their ability to predict the same external outcome criterion.

A longitudinal approach can also be used to study the possible gain in predictive power when information from different sources is combined. Schachar, Rutter, and Smith (1981) found that children who were scored high on the hyperactivity factor of both the Rutter parent and teacher questionnaire (pervasive hyperactive children) were more likely to show maladjustment 4 years later than children scored deviant by only one type of informant (situational hyperactive children). General disturbance was defined in terms of a total problem score above the clinical cutoff on either

the parent or the teacher questionnaire, and not by outcome criteria that were external to the initial measurements.

In the present longitudinal study, we investigated the ability of parent and teacher ratings of problem behaviors in children from the general population to predict outcome measures that were external to the initial measurements across a 6-year interval. Scores on the Child Behavior Checklist (CBCL; Achenbach, 1991b) and the Teacher's Report Form (TRF; Achenbach, 1991c) were used to predict outcomes. Outcome measures were based on the preceding 2 years. We then compared the predictive power of each type of informant across the 6-year time interval. We also compared the predictive power for each informant separately with that for both sources combined.

## METHOD

### *Description of Sample*

The original sample consisted of 4- to 16-year-olds drawn in 1983 (Time 1) from the Dutch province of Zuid-Holland. For the present study, only those children who were 4 to 11 years in 1983 were included. The original sample consisted of 1,311 children aged 4 to 11 years. Parents were interviewed by trained interviewers who recorded the parents' answers to each question on the CBCL (Achenbach, 1991b). All parents or guardians of children attending kindergarten or elementary school were asked for permission to have their children's teacher complete the TRF (Achenbach, 1991c).

Six years later, parents were interviewed again. Because children were assessed at 2-year time intervals, we designated this 6-year assessment as Time 4. In addition to CBCLs, outcome measures were obtained through parent interviews. If parents and children consented, the teacher was sent a TRF.

For 946 of the original 1,311 children (72.2%) we obtained usable CBCLs and TRFs at Time 1 and usable outcome variables at Time 4.

### *Instruments*

**CBCL.** The CBCL (Achenbach, 1991b) was used to obtain standardized parents' reports of children's problem behaviors. It consists of 20 competence items and 120 problem items. Only the findings from the problem section will be reported here. Parents are requested to rate the prob-

lem items on a 3-point scale based on the preceding 6 months in the following way: 0 if the item is not true of the child, 1 if the item is somewhat or sometimes true, and 2 if it is very true or often true. A total problem score is computed by summing all 0s, 1s, and 2s.

The CBCL was translated into Dutch with the help of a linguist. The good reliability and discriminative validity established by Achenbach (1991b) were confirmed for the Dutch translation (Verhulst, Akkerhuis, & Althaus, 1985; Verhulst, Berden, & Sanders-Woudstra, 1985).

A correlation of .70 has been obtained between CBCL total problem scores and problem scores based on clinical interviews with parents averaging 313 days after they completed the CBCL (Verhulst & Van der Ende, 1991b).

*TRF.* The Teacher's Report Form (Achenbach, 1991b) was used to obtain standardized teacher reports on children's behavioral/emotional problems. The TRF is the teacher version of the CBCL. The TRF has 120 problem items, including 95 of the same problem items as the CBCL. Items that were relevant to the home situation but not to the classroom situation, such as bedwetting or nightmares, were replaced by items more relevant to the school situation, such as *disrupts class discipline* and *inattentive, easily distracted*.

The teacher is asked to rate the problem items on the same 3-point scale as the CBCL, except that the TRF scores are based on the preceding 2 months. A total problem score is computed by summing all 0, 1, and 2 ratings. The TRF also contains a section on academic performance and general adaptive functioning. However, only the problem section was scored by teachers in the present study.

The TRF was translated into Dutch with the help of a linguist. The good reliability and discriminative validity established by Achenbach (1991c) were confirmed by our own studies (Verhulst & Akkerhuis, 1986; Verhulst, Akkerhuis, & Althaus, 1985). Significant associations of TRF scores with clinical psychiatric judgment further support the utility of this instrument (Verhulst, Berden, & Sanders-Woudstra, 1985b).

*Cross-Informant Syndromes.* Achenbach (1991a) has constructed eight cross-informant syndromes that can be scored on the parent, teacher, and self-report (the Youth Self-Report; YSR) versions of the CBCL. In the present study, the CBCL and TRF ratings were scored on the cross-informant syndromes. The syndromes were empirically derived via principal-components analyses of the CBCL, TRF, and YSR. Two broad-band groups of syndromes, designated as *externalizing* and *internalizing* were also used in the analyses. The internalizing group consisted of the Anxious/Depressed, Somatic Complaints, and Withdrawn syndromes. The externalizing group consisted of the Aggressive and Delinquent Behavior syndromes.

*Outcome Measures.* The parent interview at Time 4 included questions on six signs of disturbance indicating whether the child (a) had academic problems (by asking the parent whether or not the child had received any professional help, or had been referred to a professional for learning problems); (b) had behavioral problems in school (by asking the parent whether or not the child had been expelled from school as a result of misconduct, or whether or not the child had received any professional help, or had been referred to a professional for problem behavior in the school); (c) had received mental health services; (d) was judged by the parents to need professional help for behavioral/emotional problems that he/she was not receiving; (e) had shown suicidal behavior; or (f) had been in trouble with the police. Outcome variables were based on the preceding 2 years. Each outcome variable was scored 0 when this variable was not true of the child, and 1 when the outcome variable was true of the child.

## RESULTS

### *Comparison of Dropouts and Remainers*

The dropouts ( $n = 365$ ) and remainers ( $N = 946$ ) were compared with respect to age, sex, socioeconomic status (SES), and CBCL and TRF total problem scores at Time 1. The dropouts consisted of children on whom we had Time 1 CBCLs but not TRFs, as well as children who became nonresponders during the study interval. SES was measured on a six-step scale of parental occupation (Van Westerlaak, Kropman, & Collaris, 1975).

Dropouts and remainers did not differ significantly in the level of Time 1 CBCL mean total problem scores (23.5 for dropouts and 21.8 for remainers;  $t = 1.66$ ;  $df = 588$  for  $t$ -test with unequal variances; ns). However, both groups differed in their initial level of TRF mean total problem scores, with a total problem score of 22.3 for dropouts ( $n = 144$ ), and 17.7 for remainers ( $n = 943$ ; for three children it was not possible to compute a total problem score due to missing data). The difference in total problem scores was significant ( $t = 2.8$ ;  $df = 1,085$ ). Remainers and dropouts did not differ significantly in their sex and age distributions. Parents of lower SES (levels 1 to 3) were somewhat overrepresented in the group of dropouts (62% of lower SES in the group of dropouts, and 54% in the group of remainers) ( $\chi^2 = 6.73$ ;  $df = 1$ ;  $p < .01$ ). From the comparison between dropouts and remainers, it is clear that children from lower SES and with relatively high initial levels of teacher-reported problem behaviors were slightly underrepresented in the study sample. The underrepresentation of

relatively problematic and disadvantaged subjects may have slightly weakened the power of the initial problem scores to predict outcomes.

### *The Prevalence of Outcome Signs*

Table I shows the prevalence rates of the different signs of poor outcome for boys and girls separately. As can be seen, the prevalence of poor outcomes is slightly higher for boys (17.6%) than for girls (11.3%). A significant sex difference was found for police contact and academic problems as well as for any poor outcome, which were scored more frequently for boys than for girls.

### *The Additive Effect of Parent and Teacher Information*

We identified children whose Time 1 CBCL or TRF problem scores were deviant versus nondeviant. Deviance was operationally defined as a score above the 85th percentile of the cumulative frequency distribution for the internalizing, externalizing, and total problem scores. Deviant scores on the syndrome scales were defined as a *T* score  $\geq 67$ , which is approximately the 95.5th percentile. Cutoffs were determined for each sex separately and were based on the Dutch norms. For each CBCL and TRF scale, four categories were formed on the basis of Time 1 problem scores, consisting of children who scored (1) nondeviant on both the CBCL and TRF, (2) deviant on the CBCL only, (3) deviant on the TRF only, and (4) deviant on both the CBCL and TRF.

Table II shows the percentage of boys and girls in each of the four Time 1 categories who had poor outcomes on any signs of disturbance.

Table I. Percentage of Children in Each Outcome Category

	Boys ( <i>n</i> = 539)	Girls ( <i>n</i> = 567)	<i>p</i> <sup>a</sup>
Referral to mental health service	4.8	3.0	ns
Academic problems	5.6	2.8	<.05
School behavior problems	3.2	1.4	ns
Police contact	3.7	0.9	<.01
Suicidal behavior	0.9	0.9	ns
Need for help	4.6	4.2	ns
Any poor outcome	17.6	11.3	<.01

<sup>a</sup>Chi square test for differences in proportions.

Table II. Percentage of Boys and Girls with Poor Outcomes by Time 1 CBCL and TRF Categories for Each Scale<sup>a</sup>

Scale	Time 1 category												p value <sup>b</sup>				
	CBCL and TRF nondeviant				CBCL deviant				TRF deviant					CBCL and TRF deviant			
	Boys % (n)	Girls % (n)	Boys % (n)	Girls % (n)	Boys % (n)	Girls % (n)	Boys % (n)	Girls % (n)	Boys % (n)	Girls % (n)	Boys % (n)	Girls % (n)		Boys % (n)	Girls % (n)		
Withdrawn	17(411)	10(460)	25(12)	13(15)	22(18)	37(19)	20(5)	67(3)	ns	<.001							
Somatic Complaints	16(403)	11(455)	24(29)	14(22)	33(12)	40(15)	50(2)	20(5)	ns	<.01							
Anxious/Depressed	16(401)	11(447)	29(21)	18(22)	29(17)	20(25)	43(7)	67(3)	<.05	<.05							
Social Problems	17(410)	10(460)	21(14)	15(13)	22(18)	29(17)	50(4)	86(7)	ns	<.001							
Thought Problems	16(415)	11(449)	29(14)	18(33)	18(17)	18(11)	-(-)	25(4)	-	ns							
Attention Problems	16(407)	9(463)	26(19)	55(11)	25(16)	36(14)	75(4)	56(9)	<.05	<.001							
Delinquent Behavior	15(407)	10(449)	42(19)	21(24)	33(18)	33(21)	50(2)	67(3)	<.01	<.001							
Aggressive Behavior	15(409)	10(459)	44(18)	28(18)	38(16)	38(16)	-(-)	50(4)	<.01	<.001							
Internalizing	15(324)	9(352)	19(57)	13(70)	23(43)	20(51)	36(22)	33(24)	<.05	<.01							
Externalizing	13(325)	7(359)	26(54)	15(52)	23(47)	22(65)	45(20)	48(21)	<.001	<.001							
Total Problems	13(323)	7(375)	25(53)	14(50)	27(48)	19(47)	36(22)	56(25)	<.01	<.001							

<sup>a</sup>The numbers of boys and girls in each CBCL and TRF category are shown in parentheses; the percentages indicate the percentage of the numbers of subjects shown in parentheses. CBCL = Child Behavior Checklist; TRF = Teacher's Report Form.

<sup>b</sup>Fisher's Exact Test for  $r \times c$  tables (Mehta & Patel, 1983).



For each sex, differences in the proportions of children across the categories were tested by Fisher's Exact Test for  $r \times c$  tables (Mehta & Patèl, 1983) (right column). For the majority of scales, the association with later poor outcome increased if Time 1 scores were in the deviant range on both instruments. Girls had a greater probability than boys to show no signs of disturbance at followup if they scored in the nondeviant range on both the CBCL and TRF at Time 1. Except for the Thought Problems syndrome, all other scales for girls were associated with later poor outcomes if one or both instruments had syndrome scores in the deviant range. For boys, scores on five of the eight syndromes, the externalizing and internalizing scales, and the total problem score were associated with later signs of disturbance. Note that the predictive power of the TRF for girls is on the average somewhat greater than that of the CBCL. The combination of deviant CBCL and TRF scores was clearly associated with poor outcome: 56% of the girls and 36% of the boys with total problem scores in the deviant range on both instruments could be regarded as maladjusted 6 years later; this was true for 48% of the girls and 45% of the boys with externalizing scores, and for 33% of the girls and 36% of the boys with internalizing scores in the deviant range of the CBCL and TRF.

Table III shows the percentages of children in each of the poor outcome categories by Time 1 CBCL and TRF categories for the Externalizing, Internalizing, and Total Problem scales. The numbers of children in each cell for the specific cross-informant syndromes were too small to be meaningful, and are therefore not reported here. The association between deviance on the CBCL and TRF at initial assessment with reports of parents' need for help without receiving it or referral for mental health services 6 years later was stronger than associations with the other outcome measures. It was striking that the combination of parents' and teacher's reports of deviance predicted parental judgment of need for help (without receiving it) across a 6-year time span in 26% of the cases, and referral for mental health services in another 15%. Subjects could only be scored on either the variable concerning need for help or concerning referral to mental health services, but not both. For the other variables, subjects could be scored on more than one. The fact that parents of 41% of the children who were initially judged deviant by both parents and teachers reported specific signs of disturbance 6 years later clearly showed the strong predictive power of parent and teacher reports assessed in a standardized way via the CBCL and TRF.

Table III also shows that teacher information was equal to, or even somewhat better than, parent information in predicting later signs of disturbance. Externalizing and internalizing problems predicted referral for mental health services to about the same degree. Externalizing problems,

**Table III.** Percentage of Children in Each Outcome Category by Time 1 CBCL and TRF Categories for the Internalizing, Externalizing, and Total Problems Scale

	Time 1 Category				<i>p</i> <sup>b</sup>
	CBCL and TRF nondeviant % ( <i>n</i> ) <sup>a</sup>	CBCL deviant % ( <i>n</i> ) <sup>a</sup>	TRF deviant % ( <i>n</i> ) <sup>a</sup>	CBCL and TRF deviant % ( <i>n</i> ) <sup>a</sup>	
<b>Referral to mental health services</b>					
Internalizing	3(676)	6(127)	6(94)	15(46)	<.001
Externalizing	2(684)	5(106)	7(112)	20(41)	<.001
Total Problems	3(698)	4(103)	7(95)	15(47)	<.001
<b>Academic problems</b>					
Internalizing	4	2	4	7	ns
Externalizing	4	6	4	15	<.01
Total Problems	3	5	6	11	ns
<b>School behavior problems</b>					
Internalizing	2	2	—	4	ns
Externalizing	1	5	2	10	<.01
Total Problems	1	5	2	6	<.05
<b>Police contact</b>					
Internalizing	1	3	5	2	<.05
Externalizing	1	4	4	2	<.05
Total Problems	1	3	3	9	<.01
<b>Suicidal behavior</b>					
Internalizing	1	—	1	4	ns
Externalizing	1	1	2	—	ns
Total Problems	1	—	1	4	ns
<b>Need for help</b>					
Internalizing	3	5	10	17	<.001
Externalizing	3	8	9	17	<.001
Total Problems	3	4	7	26	<.001

<sup>a</sup>Numbers of subjects are given for the first three rows only because they are similar for the corresponding rows for each poor outcome variable. CBCL = Child Behavior Checklist; TRF = Teacher's Report Form.

<sup>b</sup>Fisher's Exact Test for difference in proportions.

but not internalizing problems, predicted academic problems as well as school behavior problems. There was a weak association between initial problems scored by parents and/or teachers and later trouble with the police. The additive effect of the combination of parent and teacher information was only found for total problem scores. Suicidal behavior was not significantly predicted on the basis of initial problem scores.

*The Prediction of Poor Outcomes*

Table IV shows the odds ratios and 95% confidence intervals for each CBCL and TRF problem scale indicating the strength of the association with poor outcomes. The outcome variable was dichotomized: A subject was scored 0 if there was no sign of disturbance at followup, and 1 if the subject scored on one or more of the poor outcome variables.

As can be seen from Table IV, the majority of TRF scales tended to predict poor outcomes better than the corresponding CBCL scales, although none of the differences were significant, because the confidence intervals for each pair of corresponding scales showed an overlap. Within each instrument (CBCL and TRF) Attention Problems predicted outcome best, although the odds ratio for Attention Problems did not differ significantly from the odds ratios for other scales within the same instrument.

Because the CBCL and TRF scale scores are not independent from each other, between and within the instruments, we performed stepwise logistic regressions on three different sets of predictor variables. The first

Table IV. Odds Ratios and 95% Confidence Limits of Time 1 CBCL and TRF Syndrome Scores Against Poor Outcomes<sup>a</sup>

CBCL/TRF scales	CBCL		TRF	
	Odds ratio	(95% CI)	Odds ratio	(95% CI)
Withdrawn	1.8	(0.8-4.0)	2.9	(1.5-5.6)
Somatic Complaints	1.6	(0.8-3.1)	3.5	(1.7-7.2)
Anxious Depressed	2.5	(1.4-4.8)	2.6	(1.4-4.9)
Social Problems	3.4	(1.7-6.7)	3.9	(2.1-7.3)
Thought Problems	1.7	(0.9-3.4)	1.4	(0.6-3.5)
Attention Problems	5.4	(2.9-10.1)	4.3	(2.3-8.2)
Delinquent Behavior	3.3	(1.7-6.1)	3.7	(2.0-7.1)
Aggressive Behavior	3.3	(1.7-6.2)	3.3	(1.7-6.6)
Internalizing	1.8	(1.2-2.7)	2.5	(1.6-3.8)
Externalizing	2.9	(1.9-4.4)	3.1	(2.1-4.7)
Total Problems	2.9	(1.9-4.5)	3.5	(2.3-5.3)

<sup>a</sup>Poor outcome was defined as any of the following six signs of disturbance indicating whether the child (a) had academic problems, (b) had behavioral problems in school, (c) had received mental health services, (d) was judged by the parents to need professional help for behavioral/emotional problems that he/she was not receiving, (e) had shown suicidal behavior, or (f) had been in trouble with the police. Outcome variables were based on the preceding 2 years. CI = confidence interval; CBCL = Child Behavior Checklist; TRF = Teacher's Report Form.

set included the eight cross-informant syndromes scored for the CBCL and TRF, the second set included CBCL and TRF externalizing and internalizing scores, and the third set included CBCL and TRF total problem scores. In each analysis, predictor variables also included age, sex, and SES. Table V shows the results of the three analyses. Only those predictor variables are listed that were entered and not removed in the stepwise regression procedure.

The odds ratio indicates the strength of the association with the outcome variable adjusted for the contribution of the other variables that were included. Values greater than 1.0 indicate a positive association with the outcome, while values less than 1.0 indicate a negative relation with outcome. Table V also gives the 95% confidence limits for each odds ratio.

Table V shows that two of the CBCL and three of the TRF cross-informant scales significantly predicted poor outcomes across the 6-year

Table V. Odds Ratios and 95% Confidence Intervals in Logistic Regression Analyses of Time 1 Predictor Variables Against Poor Outcome<sup>a</sup>

Variables <sup>c</sup>	Odds ratio	(95% confidence interval)
<b>I</b>		
Sex	0.61 <sup>b</sup>	(0.41-0.89)
CBCL Attention Problems	2.98	(1.49-5.98)
CBCL Delinquent Behavior	2.62	(1.32-5.19)
TRF Somatic Complaints	2.40	(1.05-5.49)
TRF Social Problems	2.28	(1.12-4.66)
TRF Delinquent Behavior	2.33	(1.15-4.73)
<b>II</b>		
Sex	0.61 <sup>b</sup>	(0.42-0.90)
CBCL Externalizing	2.41	(1.56-3.73)
TRF Internalizing	1.86	(1.17-2.95)
TRF Externalizing	2.50	(1.62-3.88)
<b>III</b>		
Sex	0.64 <sup>b</sup>	(0.44-0.93)
CBCL Total Problems	2.36	(1.53-3.65)
TRF Total Problems	2.96	(1.92-4.56)

<sup>a</sup>Poor outcome was defined as any of the following six signs of disturbance indicating whether the child (a) had academic problems, (b) had behavioral problems in school, (c) had received mental health services, (d) was judged by the parents to need professional help for behavioral/emotional problems that he/she was not receiving, (e) had shown suicidal behavior, or (f) had been in trouble with the police. Outcome variables were based on the preceding 2 years. CBCL = Child Behavior Checklist; TRF = Teacher's Report Form.

<sup>b</sup>Girls had a smaller probability of poor outcome than boys; all other odds ratios indicate a higher probability of poor outcome for children scored in the clinical range of a particular problem scale.

<sup>c</sup>Only those variables are listed that were entered and not removed in the three sets of stepwise logistic regressions.

interval. The variable in this analysis that independently predicted outcome best was the CBCL Attention Problems scale ( $OR = 2.98$ ). This means that children who were scored in the deviant range of the CBCL Attention Problems scale were nearly three times more likely to show signs of disturbance than children who were not scored deviant by their parents at initial assessment. Both the CBCL and TRF Delinquent Behavior scales significantly predicted poor outcomes. Two of the three TRF internalizing scales (Somatic Complaints and Social Problems) were also significant predictors of poor outcomes.

The second analysis showed that the CBCL and TRF externalizing scales both predicted later signs of disturbance. The TRF internalizing scale also significantly predicted poor outcomes, but to a lesser degree than the externalizing scale on both instruments. However, the difference in odds ratios for the TRF internalizing versus externalizing scales was not significant because their 95% confidence limits showed an overlap.

The third analysis showed that on both instruments the total problem score was a significant predictor, with the TRF total problem score showing a somewhat larger odds ratio than the CBCL total problem score.

The odds ratio for sex in each of the three analyses indicated that girls had a smaller probability for signs of disturbance at followup than boys. Age and SES did not significantly predict poor outcomes.

## DISCUSSION

The present study showed that, even across 6 years, CBCL and TRF scores significantly predicted signs of disturbance. These findings support the validity of both instruments. It can be argued that two of these outcome criteria, referral for mental health services and need for help, are signs of disturbance that are related to parental perception of the child's problems, and are therefore not truly external to the CBCL. However, Time 1 TRF scores obtained from teachers showed an even greater association with these variables than Time 1 CBCL scores obtained from parents.

Comparison of the predictive power of parent and teacher information showed that teacher scores were a stronger predictor of poor outcome than parent scores. Although teachers report less problems than parents about the same children (Verhulst & Akkerhuis, 1989), their reports apparently are informative with respect to later functioning. It is therefore important to include teacher information in the diagnostic assessment of children.

The combination of parent and teacher information resulted in an improvement of predictive power over that of information from one source

only. This additive effect was clearer for girls than for boys and greater for externalizing than for internalizing problems. These results showed the importance of obtaining assessments from both parents and teachers. Our data did not allow the study of the effects of adding self-reports to information from parents and teachers.

Externalizing problems predicted later signs of disturbance to a somewhat greater degree than internalizing problems. A factor possibly explaining the difference between externalizing and internalizing scores is that the outcome variables in our study usually involved the judgment of children's functioning by adults. Adults may be more alert to problems that bring a child into conflict with other people than to problems reflecting the child's inner distress. It is also possible that externalizing problems evoke environmental reactions that lead to other problems which may, in turn, promote signs of general maladjustment.

Of the eight CBCL syndromes, Attention Problems and Delinquent Behavior significantly predicted poor outcome. Delinquent Behavior was also among the TRF syndromes that significantly predicted poor outcome. In a 3-year followup of a U.S. national sample of 4- to 16-year-olds, the Attention Problems syndrome predicted academic problems, school behavior problems, and total disturbance (Stanger, Achenbach & McConaughy, 1993). Attention Problems was also a predictor of the Delinquent Behavior syndrome across the 3-year period (Stanger, McConaughy, & Achenbach, 1992). These findings also agree with reports on the stability of delinquent child behaviors (Loeber, 1982), and on the poor outcome of children with attention deficit hyperactivity disorder (Klein & Manuzza, 1991). The findings show that parent-reported delinquent behaviors and attention problems at ages 4 to 11 years predicted signs of general disturbance across 6 years. No parent-reported internalizing syndrome scales independently predicted poor outcomes.

The TRF Somatic Complaints and Social Problems syndromes also predicted poor outcomes. Apparently these problems perceived by a child's teacher are indicative of later maladjustment. Because teachers can evaluate how children interact in the classroom as well as during recess, teachers may be valuable informants on behaviors constituting the Social Problems syndrome, including acts too young, is too dependent, doesn't get along with other children, is teased, and is not liked by other children. These problems lack a clear counterpart in DSM-III-R (American Psychiatric Association, 1987) nosology. The present study showed that teacher-reported social problems significantly predicted later signs of disturbance. Interventions to improve the social skills of children with poor peer relationships may help in bending the development course in a more favorable direction.

The present study investigated the prediction of signs of disturbance that were largely related to the perception of problems by adults in the child's environment. A next step is to investigate the relation between early predictors and signs of disturbance that are related to the subject's own perceptions. Followups of children into young adulthood will make it possible to test the predictive power of parent and teacher assessments against signs of disturbance including young adults' own decisions to seek help, academic and professional career, and the ability to form stable relationships.

## REFERENCES

- Achenbach, T. M. (1982). *Developmental psychopathology*. New York: John Wiley and Sons.
- Achenbach, T. M. (1985). *Assessment and taxonomy of child and adolescent psychopathology*. Beverly Hills, CA: Sage.
- Achenbach, T. M. (1991a). *Integrative guide for the 1991 CBCL/4-18, YSR, and TRF profiles*. Burlington: University of Vermont, Department of Psychiatry.
- Achenbach, T. M. (1991b). *Manual for the CBCL/4-18 and 1991 profile*. Burlington: University of Vermont, Department of Psychiatry.
- Achenbach, T. M. (1991c). *Manual for the TRF and 1991 profile*. Burlington: University of Vermont, Department of Psychiatry.
- Achenbach, T. M., McConaughy, S. H., & Howell, C. T. (1987). Child/adolescent behavioral and emotional problems: Implication of cross-informant correlations for situational specificity. *Psychological Bulletin*, *101*, 213-332.
- American Psychiatric Association. (1987). *Diagnostic and statistical manual of mental disorders* (3rd ed., rev.). Washington, DC: Author.
- Bird, H., Gould, M. S., & Staghezza, B. (1992). Aggregating data from multiple informants in child psychiatry epidemiological research. *Journal of the American Academy of Child and Adolescent Psychiatry*, *31*, 78-85.
- Klein, R., & Manuzza, S. (1991). Long-term outcome of hyperactive children: A review. *Journal of the American Academy of Child and Adolescent Psychiatry*, *30*, 383-387.
- Loeber, R. (1982). The stability of antisocial and delinquent child behavior: A review. *Child Development*, *53*, 1431-1446.
- Loeber, R., Green, S. M., & Lahey, B. B. (1990). Mental health professionals' perception of the utility of children, mothers and teachers as informants on childhood psychopathology. *Journal of Clinical Child Psychology*, *19*, 136-143.
- Loeber, R., Green, S. M., Lahey, B. B., & Stouthamer-Loeber, M. (1990). Optimal informants on childhood disruptive behaviors. *Development and Psychopathology*, *1*, 317-337.
- McConaughy, S. H., Stanger, C., & Achenbach, T. M. (1992). Three-year course of behavioral/emotional problems in a national sample of 4- to 16-year-olds: I. Agreement among informants. *Journal of the American Academy of Child and Adolescent Psychiatry*, *31*, 932-940.
- Mehta, C. R., & Patel, N. R. (1983). A network algorithm for performing Fisher's Exact Test in  $r \times c$  contingency tables. *Journal of the American Statistical Association*, *78*, 427-434.
- Reich, W., & Earls, F. (1987). Rules for making psychiatric diagnoses in children on the basis of multiple information: Preliminary strategies. *Journal of Abnormal Child Psychology*, *15*, 601-616.
- Schachar, R., Rutter, M., & Smith, A. (1981). The characteristics of situationally and pervasively hyperactive children: Implications for syndrome definition. *Journal of Child Psychology and Psychiatry*, *22*, 375-392.

- Stanger, C., Achenbach, T. M., & McConaughy, S. H. (1993). Three-year course of behavioral/emotional problems in a national sample of 4- to 16-year-olds: III. Predictors of signs of disturbance. *Journal of Consulting and Clinical Psychology, 61*, 839-848.
- Stanger, C., McConaughy, S. H., & Achenbach, T. M. (1992). Three-year course of behavioral/emotional problems in a national sample of 4- to 16-year-olds: II. Predictors of syndromes. *Journal of the American Academy of Child and Adolescent Psychiatry, 31*, 941-950.
- Van Westerlaak, J. M., Kropman, J. A., & Collaris, J. W. M. (1975). *Beroepenklapper* [Index of occupations]. Nijmegen, Holland: Instituut voor Toegepaste Sociologie.
- Verhulst, F. C., & Akkerhuis, G. W. (1986). Mental health in Dutch children: III. Behavioral-emotional problems reported by teachers of children aged 4-12. *Acta Psychiatrica Scandinavica* (Suppl. 330), 74.
- Verhulst, F. C., & Akkerhuis, G. W. (1989). Agreement between parents' and teachers' ratings of behavioral/emotional problems of children aged 4-12. *Journal of Child Psychology and Psychiatry, 30*, 123-136.
- Verhulst, F. C., Akkerhuis, G. W., & Althaus, M. (1985). Mental health in Dutch children: I. A cross cultural comparison. *Acta Psychiatrica Scandinavica* (Suppl. 323), 72.
- Verhulst, F. C., Berden, G. F. M., & Sanders-Woudstra, J. A. R. (1985). Mental health in Dutch children: II. Prevalence of psychiatric disorder and relationships between measures. *Acta Psychiatrica Scandinavica* (Suppl. 324), 72.
- Verhulst, F. C., Koot, J. M., & Berden, G. F. M. J. (1990). Four-year follow-up of an epidemiological sample. *Journal of the American Academy of Child and Adolescent Psychiatry, 29*, 440-448.
- Verhulst, F. C., & Van der Ende, J. (1991a). Four-year follow-up of teacher-reported problem behaviors. *Psychological Medicine, 21*, 965-977.
- Verhulst, F. C., & Van der Ende, J. (1991b). Assessment of child psychopathology: Relationships between different methods, different informants and clinical judgment of severity. *Acta Psychiatrica Scandinavica, 84*, 155-159.
- Verhulst, F. C., & Van der Ende, J. (1992a). Six-year developmental course of internalizing and externalizing problem behaviors. *Journal of the American Academy of Child and Adolescent Psychiatry, 31*, 924-931.
- Verhulst, F. C., & Van der Ende, J. (1992b). Six-year stability of parent reported problem behavior in an epidemiological sample. *Journal of Abnormal Child Psychology, 20*, 595-610.