

The Eight-Year Stability of Problem Behavior in an Epidemiologic Sample

FRANK C. VERHULST AND JAN VAN DER ENDE

Department of Child and Adolescent Psychiatry/Erasmus University Rotterdam, The Netherlands

ABSTRACT

This study assessed the 8-y stability and change of parent-reported problem behaviors in 791 children from the general population assessed at 2-y intervals with the same standardized procedure, the child behavior checklist (CBCL). Problem behaviors were rather stable across even the relatively long time interval of 8 y. The 8-y stability coefficient for total problem score was 0.48, with somewhat higher stability for externalizing than internalizing problems. Despite stability, problem behaviors also changed considerably across time. These changes do not only reflect the increase in chronologic age. The present study

showed that there are changes in children's problem levels that result from historical time influences. These secular trends seemed of even greater importance than aging effects, and reflect a slight increase of the norm of children's problem behaviors. (*Pediatr Res* 38: 612-617, 1995)

Abbreviations

CBCL, child behavior checklist
SES, socioeconomic status

To gain insight into which problems tend to persist and which problems tend to improve with increasing age of the child, we need to study the longitudinal course of these problems. (For brevity "child" will be used throughout the text to include "adolescent"). However, little is known about the development of child psychopathology. A number of factors limit the conclusions drawn from existing longitudinal studies (1), including: 1) the use of clinical samples, which may show biases as a result of factors involved in the referral process; 2) the use of different assessment and/or diagnostic procedures at different times of assessment; and 3) the lack of empirical basis for many diagnostic categories of child psychopathology.

To overcome a number of obstacles mentioned above, we have traced an epidemiologic sample of children across an 8-y time period, using the same standardized assessment procedure at each assessment. The present study was preceded by studies investigating the developmental course of problem behaviors across 2- to 6-y time intervals in Dutch children from the same general population sample (2-6). These studies revealed considerable stability of parent and teacher reported problems, as well as the ability of parent and teacher information to predict maladjustment over time. Of the 92 children aged 4-11 y who could be regarded deviant according to their parents' scores at

time 1, 33% could still be regarded deviant 6 y later. In contrast, of the 450 well functioning children (those who were scored below the 50th percentile of the frequency distribution of total problem scores), only 2% were scored in the deviant range 6 y later.

To study the 8-y developmental course of children's problem behaviors in the present study, we assessed parent-reported problems in a community sample of children with a wide age range (4-10 y), at 2-y intervals. In interpreting differences in problem scores across time, we need to distinguish cohort, period, and aging effects (7-10). A cohort is defined as a group of individuals who share the same event (*e.g.* birth) during the same time period (often, 1 y). Cohort effects, sometimes called generational effects, are the result of the membership of a particular cohort rather than another cohort. Cohort differences result from factors before the assessment. Aging effects refer to changes in individuals that occur with age, for example the increase in cognitive abilities or motor skills with increasing age in children. Aging effects reflect the ontogenetic component of change in developmental research. Cross-sectional studies confound cohort and aging effects. For example, in cross-sectional data, 4 y olds may differ from 14 y olds both as the result of differences in cohort composition (*e.g.* because of genetic factors, or deaths), and as a result of age.

Period effects refer to influences specific to a particular time period, for example economic recession, or changes in the educational system. Longitudinal studies of single cohorts confound age and period effects. For example, in longitudinal data, a child at age 4 y may differ from the same child at age

Received November 17, 1994; accepted March 20, 1995.

Manuscript dedicated to Professor H.K.A. Visser in honor of his retirement.

Correspondence: 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.

Support in part by the Sophia Foundation for Medical Research and the Dutch National Fund for Mental Health (Nationaal Fonds Geestelijke Volksgezondheid).

14 y, both as a result of the effects of changes across the time period, and as a result of effects associated with aging.

In the present study, by employing a combination of the longitudinal and cross-sectional approach in a sample of children with a broad age range, we tried to disentangle aging effects from cohort and period effects. Using this approach, we can test whether the same changes with age can be observed in different cohorts studied in different time periods.

METHODS

Population and procedure. The original sample consisted of 13 birth cohorts of children who were 4–16 y old at initial assessment in 1983 (time 1). Children were randomly selected from municipal population registers from the Dutch province of Zuid-Holland. (For details on sample selection, see Ref. 11.) Parents (in the majority, mothers) were interviewed in 1983, and reassessed at four 2-y intervals in 1985, 1987, 1989, and 1991, for a total of five measurement points. At times 1, 3, 4, and 5, parents were interviewed at home by trained interviewers who recorded the parents' answers to each question on the CBCL (12), a questionnaire designed to obtain standardized parental reports of children's problem behaviors. At time 2, the CBCL was mailed to the parents. Nonresponders were telephoned. For the present study, only children from the age cohorts 4–10 y in 1983 were included, because these were the subjects who remained within the age limits of the CBCL across the 8 y of the study. Of the 1149 children from the age cohorts 4–10 y who participated at time 1, 791 (68.8%) were assessed at all five points.

Instruments. The CBCL (12) was used to obtain standardized parents' reports of children's problem behaviors. It consists of 20 competence items and 120 problem items. The findings from only the problem section will be reported here. Parents are requested to rate the problem items on a 3-point scale based on the preceding 6 mo 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 higher the score, the more problematic is the child's behavior. The CBCL was translated into Dutch with the help of a linguist. The good reliability and discriminative validity established by Achenbach (12) were confirmed for the Dutch translation (11, 13). A correlation of 0.70 has been obtained between CBCL total problem scores and problem scores based on clinical interviews with parents averaging 313 d after they completed the CBCL (14).

Achenbach (15) has empirically constructed eight cross-informant syndromes that can be scored on the parent, teacher, and self-report versions of the CBCL. Two broad-band groups of syndromes, designated as "externalizing" and "internalizing" were also used in the present study. Externalizing problems reflect conflicts with other people and their expectations of the child, whereas internalizing problems reflect internal distress. The internalizing group consists of the Anxious/Depressed, Somatic Complaints, and Withdrawn syndromes. The externalizing group consists of the Aggressive and Delinquent Behavior syndromes.

RESULTS

Comparison of dropouts and remainers. The dropouts ($n = 358$) and remainers ($n = 791$) were compared with respect to age, sex, SES, and CBCL total problem scores at time 1. SES was measured on a 6-step scale of parental occupation (16). Dropouts and remainers did not differ significantly in the level of time 1 CBCL mean total problem scores (24.1 for dropouts and 22.2 for remainers; $t = 1.67$; $df = 583$ for t test with unequal variances; $p < 0.05$, one-tailed). Remainders and dropouts did not differ significantly in their sex and age distributions. However, dropouts had significantly lower parental occupational levels at time 1 than remainers (mean SES was 3.3 for dropouts, and 3.6 for remainers; $t = -3.53$; $df = 1143$; $p < 0.001$). In interpreting our results, it should be realized that subjects who were socioeconomically slightly disadvantaged and who were slightly more problematic were somewhat underrepresented.

Changes in mean scores. Figure 1 depicts the mean total problem scores at each time of measurement by cohort. For clarity of presentation, we combined the 7 cohorts into 2 groupings, 4–6 and 7–10 y. Table 1 shows the mean problem scores for each follow-up interval, and for both sexes and age groups 4–6 and 7–10 separately.

To test cohort, time of measurement, and sex effects, we performed 2 (cohorts 4–6, and 7–10) \times 5 (time) \times 2 (sex) repeated measures analyses of variance on the CBCL mean scale scores as shown in Table 1, with time of measurement as the within subjects factor, and cohort and sex as the between subjects factors. Table 2 shows the results of the 11 analyses of variance.

Inspection of Table 2 shows that, in general, significant effects of time outnumber those of cohort (9 versus 4 significant effects). This may be due to the fact that the cohorts span 6 y (ages 4–10), and the follow-up period 8 y. Another explanation may be that period effects are more prominent than developmental aging effects.

Total problem score. For total problem score, we found no significant cohort effect. The within-subjects comparisons showed a significant time of measurement effect, with scores dropping from time 1 to time 2, followed by a much smaller decrease at each subsequent assessment. There was a signifi-

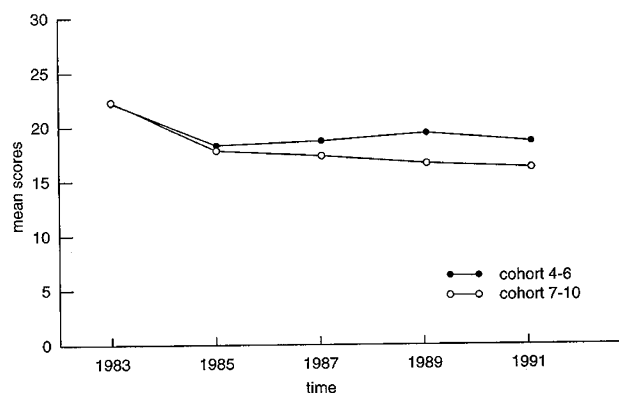


Figure 1. Mean total problem scores at the 5 points of assessment for cohorts 4–6 y and 7–10 y.

Table 1. Mean problem scores for 2-, 4-, 6-, and 8-y interval for CBCL scales by sex and ages 4–6 and 7–10 y

Problem	Time 1		Time 2				Time 3				Time 4				Time 5					
	Boys		Girls		Boys		Girls		Boys		Girls		Boys		Girls		Boys		Girls	
	4–6	7–10	4–6	7–10	4–6	7–10	4–6	7–10	4–6	7–10	4–6	7–10	4–6	7–10	4–6	7–10	4–6	7–10	4–6	7–10
Withdrawn	1.5	2.1	1.9	1.8	1.4	1.6	1.6	1.7	1.7	1.9	1.7	2.0	1.8	2.1	1.9	2.0	1.9	2.2	2.0	2.3
Somatic complaints	0.7	0.6	0.6	0.9	0.6	0.7	0.6	0.9	0.8	0.8	0.6	0.9	0.7	0.7	0.9	1.0	0.8	0.7	0.9	1.1
Anxious/depressed	1.5	3.0	2.0	2.7	1.7	2.7	1.9	2.0	2.1	2.2	2.2	2.4	2.4	2.1	2.6	2.4	2.3	1.9	2.7	2.8
Social problems	1.4	1.7	1.2	1.3	1.2	1.6	1.1	1.2	1.2	1.4	1.2	1.1	1.4	1.2	1.3	0.9	1.3	1.1	1.3	0.9
Thought problems	0.4	0.3	0.3	0.2	0.2	0.3	0.2	0.2	0.2	0.3	0.2	0.2	0.2	0.2	0.3	0.2	0.2	0.2	0.2	0.3
Attention problems	2.9	4.0	2.5	2.8	3.1	3.7	2.2	2.4	3.5	3.7	2.6	2.5	3.7	3.5	2.7	2.2	3.6	3.5	2.7	2.3
Delinquent behavior	1.5	1.5	1.1	1.0	1.4	1.2	1.0	0.7	1.2	1.1	0.8	0.7	1.3	1.5	0.8	0.7	1.3	1.6	1.0	0.9
Aggressive behavior	8.8	8.6	6.9	5.1	6.9	6.4	5.3	4.1	6.4	5.8	4.8	3.9	6.8	5.6	5.1	3.7	6.3	4.7	4.8	3.6
Internalizing	3.7	5.7	4.5	5.4	3.6	5.1	4.1	4.5	4.6	4.9	4.4	5.2	4.8	4.8	5.3	5.3	4.9	4.7	5.5	6.0
Externalizing	10.3	10.2	8.1	6.1	8.3	7.6	6.2	4.8	7.6	6.9	5.6	4.6	8.1	7.0	5.8	4.4	7.5	6.3	5.8	4.5
Total problems	23.1	25.5	21.3	19.2	19.6	20.4	17.1	15.4	20.3	19.2	17.2	15.7	20.8	18.4	18.1	14.9	19.6	16.9	17.5	15.6

Table 2. Summary of analyses of variance (F values) for CBCL scales

Source	df	Syndromes										Total problems
		Withdrawn	Somatic complaints	Anxious/depressed	Social problems	Thought problems	Attention problems	Delinquent behavior	Aggressive behavior	Internalizing	Externalizing	
Between subjects' effects												
Error	787											
Cohort (C)	1	4.6*	2.0	3.2	0.1	0.0	0.4	0.4	12.3***	4.9*	9.2**	2.4
Sex (S)	1	0.5	3.4	1.0	2.9	1.1	38.0***	31.6***	40.6***	1.6	43.4***	14.0***
C × S	1	0.8	3.5	0.4	2.3	0.2	1.3	1.0	0.6	0.1	0.8	0.8
Within subjects' effects												
Error	3148											
Time (T)	4	13.5***	2.6*	3.6**	2.2	2.7*	1.5	11.9***	53.9***	8.7***	44.6***	32.5***
C × T	4	0.1	0.4	10.3***	8.5***	1.1	6.4***	2.0	1.4	4.6**	0.5	2.6*
S × T	4	0.1	3.0*	4.2**	0.3	1.2	0.7	1.9	3.7**	2.8*	2.9*	1.5
C × T × S	4	2.0	1.1	3.3*	0.3	0.9	0.5	0.9	1.8	2.5*	1.2	1.8
T × Ss within												

Note: F and df values are derived from the multivariate Pillai's trace test statistic.

- * $p < 0.05$.
- ** $p < 0.01$.
- *** $p < 0.001$.

cant interaction between time and cohort. The mean total problem scores show a more rapid decline in scores for the older (7–10 y) than for the younger (4–6 y) cohorts. These findings indicate that period effects have exerted their influence on parent reported problem scores across the time frame of our study.

Between subjects comparisons revealed a significant sex effect for total problem score, reflecting higher total problem scores for boys over girls. There were no significant interactions between sex and cohort, and between sex and time of measurement.

Externalizing. Externalizing scores showed a significant time of measurement effect, reflecting a decline in mean scores across time. A significant time of measurement × sex interaction reflected the more rapid decline in scores for boys *versus* girls. There was no significant time of measurement × cohort interaction. Between subjects comparisons revealed a significant cohort and sex effect, but no significant sex × cohort effect, with higher scores for younger than for older subjects, and higher scores for boys over girls. The results showed that the decrease in externalizing scores can be attributed for the greatest part to aging effects. The developmental decline in aggressive behaviors is possibly related to increasing abilities of children to control their behavior.

Internalizing. Internalizing scores showed a significant time of measurement effect with scores declining from time 1 to time 2, followed by a slight increase from time 2 to time 5. There were significant time of measurement × sex, and time of measurement × cohort two-way interactions, as well as a three-way interaction of time × sex × cohort. Inspection of the mean scores revealed that the two-way interactions reflected a more rapid increase for girls over boys, and a more rapid increase for the younger over the older cohorts with increasing time of measurement. The three-way interaction reflected the increase in mean scores with time for girls and younger boys, whereas older boys showed a decrease in scores with time. Between subjects comparisons revealed no significant sex effect, and a significant cohort effect. Inspection of mean scores revealed higher scores for older *versus* younger cohorts. The results reflect the influence of secular change on internalizing scores, especially for boys.

Specific syndromes. For six of the eight syndromes, a significant time of measurement effect was found. Inspection of the mean scores revealed that for all effects the means showed a drop in scores from time 1 to time 2. Scores for the Withdrawn, Somatic Complaints and the Anxious/Depressed syndromes increased from time 2 onward. Scores for the Thought Problems syndrome increased slightly after time 2.

The same was found for the Delinquent Behavior syndrome, whereas scores for the Aggressive Behavior syndrome showed a clear decline across the five assessments.

The only two syndromes for which a significant between subjects cohort effect was found were the Withdrawn and Aggressive Behavior syndromes. For each of these two syndromes, the cohort effects were in line with the significant within subjects time effects because for neither of the two syndromes a significant cohort \times time interaction was found. The increase for the Withdrawn syndrome and the decrease for the Aggressive Behavior syndrome can be regarded as developmental aging effects.

Between subjects sex effects were found for the Attention Problems, Delinquent Behavior and Aggressive Behavior scales, with higher scores for boys over girls. For none of the scales we found a significant sex \times cohort interaction.

Significant cohort \times time interactions were found for the Anxious/Depressed, Social Problems, and Attention Problems syndromes. Inspection of the mean scores showed a slight increase in scores for younger and a slight decrease for older children. For none of these three syndromes we found a significant cohort effect. Period effects rather than aging effects may be responsible for the differences in scores on these syndromes.

Significant sex \times time interactions for Anxious/Depressed and Somatic Complaints reflected a more rapid increase of scores with time for girls over boys. A significant sex \times time interaction for Aggressive Behavior was found. Mean Aggressive Behavior scores tended to show a more rapid decline for boys over girls, with boys showing higher overall levels of aggressive behavior at each assessment.

Stability. Pearson correlations (stability coefficients) between scores at different times of assessment gives us an

indication to what extent the subjects' scores in the sample keep their relative position or ranking across time, irrespective of possible changes in mean scores. Table 3 gives the stability coefficients for each CBCL scale and for the 2-, 4-, 6-, and 8-y intervals by age group and sex.

Figure 2 shows the pattern of stability coefficients for the total sample. All correlations depicted in Figure 2 were significant. Stability coefficients decreased with increasing time interval, although a stability coefficient of 0.48 for the CBCL total problem score across the rather long time interval of 8 y can still be regarded considerable. The pattern of differences in stabilities between the different scales shows consistency, with the highest stabilities for the Aggressive Behavior syndrome for each of the four time intervals. The 8-y stability for Externalizing scores (0.50) was somewhat greater than that for Internalizing scores (0.39). Table 3 shows that the majority of stability coefficients were somewhat greater for older than for younger children in our sample. Even across the 8-y interval there was considerable stability, with stability coefficients for the total sample >0.50 for one scale (Aggressive Behavior), between 0.30 and 0.50 for eight scales, between 0.20 and 0.30 for one scale (Somatic Complaints), and only one <0.20 for Thought Problems ($r = 0.16$). Note that the highest stability was found for the Aggressive Behavior scale, whereas this scale showed the largest decline in scores across time. Apparently the most aggressive children at initial assessment tended to be the most aggressive children at later assessments irrespective of the decrease in overall levels of aggression.

Individual course of problem behaviors. To determine the individual course of problem behaviors, we looked at the percentage of children who could be regarded deviant at Time 1 and who were still deviant at follow-up. Table 4 shows the results. We took the 85th percentile of the cumulative fre-

Table 3. Correlations for 2-, 4-, 6-, and 8-y interval for CBCL scales by sex and ages 4-6 and 7-10 y

Problem	Years															
	2				4				6				8			
	Boys		Girls		Boys		Girls		Boys		Girls		Boys		Girls	
	4-6	7-10	4-6	7-10	4-6	7-10	4-6	7-10	4-6	7-10	4-6	7-10	4-6	7-10	4-6	7-10
Withdrawn	0.53	0.68	0.57	0.59	0.37	0.43	0.55	0.48	0.45	0.49	0.52	0.41	0.35	0.35	0.34	0.39
Somatic complaints	0.33	0.46	0.32	0.45	0.33	0.39	0.25	0.45	0.31	0.41	0.21	0.33	(0.11)	0.32	0.13	0.38
Anxious/depressed	0.62	0.67	0.41	0.59	0.53	0.56	0.39	0.50	0.47	0.39	0.35	0.54	0.36	0.38	0.23	0.52
Social problems	0.42	0.65	0.41	0.44	0.26	0.49	0.42	0.34	0.21	0.43	0.36	0.38	0.15	0.41	0.36	0.37
Thought problems	0.44	(0.11)	0.30	(0.12)	0.23	0.27	0.24	0.17	0.22	0.25	0.17	(0.09)	0.15	0.25	(0.04)	0.18
Attention problems	0.53	0.65	0.56	0.61	0.50	0.63	0.57	0.60	0.43	0.40	0.49	0.50	0.28	0.44	0.38	0.46
Delinquent behavior	0.42	0.42	0.42	0.51	0.38	0.51	0.49	0.42	0.45	0.45	0.30	0.29	0.33	0.43	0.20	0.32
Aggressive behavior	0.65	0.72	0.60	0.69	0.60	0.66	0.65	0.64	0.52	0.48	0.51	0.62	0.48	0.53	0.39	0.56
Internalizing	0.61	0.75	0.50	0.63	0.46	0.57	0.49	0.57	0.52	0.46	0.44	0.55	0.36	0.37	0.25	0.53
Externalizing	0.63	0.69	0.60	0.69	0.57	0.65	0.65	0.63	0.52	0.49	0.49	0.61	0.46	0.54	0.37	0.54
Total problems	0.65	0.76	0.58	0.67	0.58	0.68	0.64	0.65	0.57	0.47	0.54	0.64	0.42	0.53	0.40	0.57

Note: All correlations significant at $p < 0.05$, except those in parentheses.

DISCUSSION

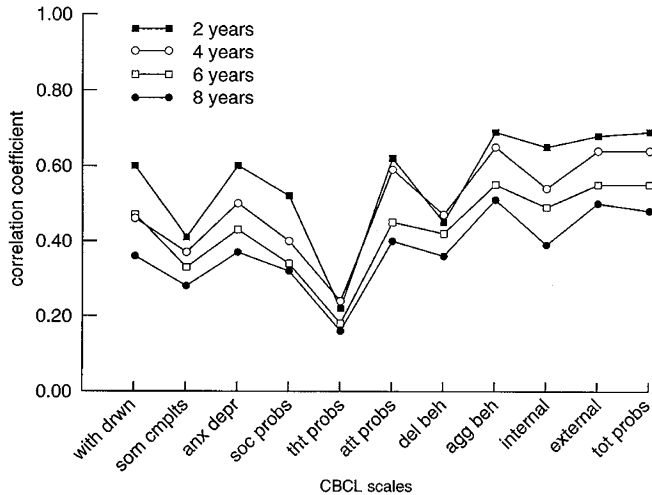


Figure 2. The 2-, 4-, 6-, and 8-y stability coefficients of CBCL scales.

Table 4. Percentage of children scoring in the deviant range of the CBCL scale scores at time 1, who still scored in the deviant range at follow-up, and odds ratios (OR) indicating the risk for deviance at follow-up given a deviant score at time 1 relative to the risk for deviance given a nondeviant score at time 1

Problem	Years							
	2		4		6		8	
	%	(OR)	%	(OR)	%	(OR)	%	(OR)
Withdrawn	41.4	(40.7)	24.1	(13.9)	37.9	(20.6)	20.7	(7.7)
Somatic complaints	18.9	(4.7)	24.5	(6.9)	18.9	(6.1)	22.6	(6.5)
Anxious/depressed	43.6	(29.8)	30.8	(11.9)	30.8	(11.1)	28.2	(8.3)
Social problems	37.0	(19.0)	22.2	(6.8)	18.5	(5.2)	25.9	(7.8)
Thought problems	7.9	(3.5)	10.5	(3.4)	15.8	(4.9)	15.8	(4.2)
Attention problems	41.7	(25.0)	33.3	(14.0)	22.2	(8.3)	19.4	(7.4)
Delinquent behavior	33.3	(21.5)	31.0	(19.3)	19.0	(4.5)	23.8	(6.0)
Aggressive behavior	28.2	(41.8)	17.9	(12.4)	17.9	(11.5)	20.5	(8.2)
Internalizing	50.4	(14.9)	41.4	(7.0)	42.1	(7.2)	33.8	(3.2)
Externalizing	44.3	(12.3)	41.2	(8.4)	41.2	(5.5)	45.0	(5.4)
Total problems	43.0	(13.2)	46.5	(13.2)	42.3	(7.2)	38.7	(5.1)

quency distribution of total problem scores of normative children as the cutoff above which children were regarded deviant. For the internalizing and externalizing scores we took the 85th percentile, and for the eight specific syndrome scales we took the 95th percentile as the cutoff. These scores are recommended by Achenbach (12) as the bottom of the deviant or clinical range.

We also wanted to quantify the relative risk of deviance at follow-up for children who could be regarded deviant initially at time 1, relative to the risk of being deviant at follow-up given a nondeviant score at time 1. Table 4 also gives the odds ratios as a measure of risk for later deviance given a score in the deviant range at time 1.

As can be seen from Table 4, nearly 39% of the deviant children at time 1 could be regarded deviant 8 y later. Internalizing problems were somewhat less persistent than externalizing problems, although the odds ratios across the 8-y interval for both syndromes did not differ significantly. Of the specific syndromes Anxious/Depressed and Aggressive Behavior were among the most persistent.

In the present study, the stability and change of parent-reported problems in children aged 4–10 y from the general population was assessed across an 8-y period with 2-y intervals and using the same standardized assessment procedure. In this way we were able to investigate the developmental trajectories of children who were of different ages at the start of the study. By employing both a cross-sectional and longitudinal approach we tried to determine whether changes in children’s problem behavior over time could be attributed to chronologic age, or to the influences of historical time to which the children have been exposed. Little is known so far to which extent the development of problem behaviors can be regarded as an invariant function of increasing age (for example a decrease in children’s overt physical aggression with increasing abilities to control behavior as a result of genetically determined maturation), or as a function of historical trends (for example an increase in children’s aggression as a result of decreasing parental control and loosening of family ties).

From the present study it is clear that time effects were more prominent than cohort effects. For instance, for total problem scores no significant cohort effect was found, whereas there was a significant time of measurement effect as well as an interaction between time of measurement and cohort. Older children showed a different trajectory than younger children, with scores declining more rapidly in the older than younger cohorts. In other words, the normative level of children’s problem behaviors has been shifted upward across the 8-y study interval from 1983 to 1991. This “secular” change in CBCL problem scores is confirmed by the historical increase in CBCL problem scores in American children from 1976 and 1989 of the same age (17). Except for total problem score in the present study, this secular trend was also found for the CBCL scales Internalizing, Anxious/Depressed, and Social Problems. The present study did not allow to test possible causes for this trend. Realizing that this explanation is only very speculative, it may be that societal changes, such as an increasing pressure on children toward greater independence from their parents, and an increasing emphasis on achievement and competition, may increase feelings of anxiety and depression and may put peer relationships under pressure.

Other areas of problem behavior seemed less influenced by this historical trend. The increase in scores with time on the Withdrawn scale and the decrease on the Aggressive Behavior scale followed similar trajectories for children from different cohorts and in different time periods. For these scales, the cross-sectional results lead to the same conclusions as the longitudinal results. The changes in scores across time reflected developmental aging effects.

For the externalizing scales Aggressive Behavior and Delinquent Behavior, and for the Attention Problems scale, we found sex differences reflecting higher scores for boys over girls. The interaction between sex and time for Aggressive Behavior scores, reflecting a more rapid decrease for boys than for girls, in combination with the absence of a significant sex and cohort, resulted in a less pronounced sex difference in the level of aggressive behaviors across the duration of the study

interval. Although speculative, it may be that changes in sex-role boundaries are responsible for this decrease in sex difference with time.

As expected, stability coefficients decreased steadily with time. However, considering the long time interval of 8 y it is surprising that for the majority of CBCL scales, the stabilities were greater than 0.30. The highest stabilities were found for the Aggressive Behavior scale, for the Externalizing scale and for the total problem scale, with correlations ranging from 0.48 to 0.51. Of the specific syndrome scales, Attention Problems and Aggressive Behavior were the problems with the highest levels of persistence. Apparently, many children with aggressive behaviors or attention problems will still show these problems to a considerable degree compared with children from the same age. Stability coefficients were somewhat greater for older than for younger children.

From a categorical point of view, a considerable proportion of children who could be regarded deviant at time 1 in different domains of psychopathology could still be regarded deviant 8 y later. Nearly 39% of the children who were maladjusted at time 1 were still so 8 y later. Odds ratios indicated that the probability of showing deviant scores at follow-up for children who could be regarded deviant at time 1 was, depending on the follow-up interval, between 5 and 13 times the probability of deviance at follow-up given a nondeviant score initially. The highest odds ratios were for Withdrawn and Anxious/depressed as well as for Aggressive Behavior. These data stress the need for adequate interventions for children who show significant levels of problems.

In conclusion, the present study showed that, although there is persistence of children's problem behaviors, these behaviors also show considerable change. However, changes in problem behaviors across time do not only reflect the increase in

chronologic age. This study showed that there seem to be changes in children's problem levels that result from historical time influences. These secular trends reflect a slight increase of the norm of children's problem behaviors.

REFERENCES

1. Verhulst FC, Koot JM 1991 Longitudinal research in child and adolescent psychopathology (special section). *J Am Acad Child Adolesc Psychiatry* 30:361-368
2. Verhulst FC, Althaus M 1988 Persistence and change in behavioral/emotional problems reported by parents of children aged 4:14: an epidemiological study. *Acta Psychiatr Scand* 77(suppl):339
3. Verhulst FC, Koot JM, Berden GFMG 1990 Four-year follow-up of an epidemiological sample. *J Am Acad Child Adolesc Psychiatry* 29:440-448
4. Verhulst FC, Van der Ende J 1991 Four-year follow-up of teacher-reported problem behaviours. *Psychol Med* 21:965-977
5. Verhulst FC, Van der Ende J 1992 Six-year developmental course of internalizing and externalizing problem behaviors. *J Am Acad Child Adolesc Psychiatry* 31:924-931
6. Verhulst FC, Koot HM, Van der Ende J 1994 Differential predictive value of parents' and teachers' reports: a longitudinal study. *J Abn Child Psychol* 22:531-546
7. Farrington DP 1991 Longitudinal research strategies: advantages, problems and prospects. *J Am Acad Child Adolesc Psychiatry* 30:369-374
8. Loeber R, Farrington DP 1995 Longitudinal approaches in epidemiological research of conduct problems. In Verhulst FC, Koot HM (eds) *The Epidemiology of Child and Adolescent Psychopathology*. Oxford University Press, London
9. Nesselroade JR, Baltes PB 1974 Adolescent personality development and historical change: 1970:1972. *Monogr Soc Res Child Dev* 39:1-80
10. Stanger C, Verhulst FC 1995 Accelerated longitudinal designs. In Verhulst FC, Koot HM (eds) *The Epidemiology of Child and Adolescent Psychopathology*. Oxford University Press, London
11. Verhulst FC, Akkerhuis GW, Althaus M 1985 Mental health in Dutch children. I. A cross cultural comparison. *Acta Psychiatr Scand* 72(suppl):323
12. Achenbach TM 1991 Manual for the Child Behavior Checklist/4-18 and 1991 Profile. University of Vermont, Department of Psychiatry, Burlington, VT
13. Verhulst FC, Berden G, Sanders-Woudstra JAR 1985 Mental health in Dutch children: II. prevalence of psychiatric disorders and relationships between measures. *Acta Psychiatr Scand* 72(suppl):324
14. Verhulst FC, Van der Ende J 1991 Assessment of child psychopathology: relationships between different methods, different informants and clinical judgment of severity. *Acta Psychiatr Scand* 84:155-159
15. Achenbach TM 1991 Integrative Guide to the 1991 CBCL/4-18, YSR and TRF Profiles. University of Vermont, Department of Psychiatry, Burlington, VT
16. Van Westerlaak JM, Kropman JA, Collaris JWM 1975 Beroepenklapper (Index of Occupations). Instituut voor Toegepaste Sociologie, Nijmegen, The Netherlands
17. Achenbach TM, Howell CT 1993 Are American children's problems getting worse? A 13-year comparison. *J Am Acad Child Adolesc Psychiatry* 32:1145-1154