

F. Almqvist  
K. Kumpulainen  
K. Ikäheimo  
S.-L. Linna  
I. Henttonen  
E. Huikko  
E. Tuompo-Johansson  
E. Aronen  
K. Puura  
J. Piha  
T. Tamminen  
E. Räsänen  
I. Moilanen

---

Professor F. Almqvist (✉) · E. Huikko  
E. Tuompo-Johansson · E. Aronen  
Department of Child Psychiatry  
University of Helsinki  
Lastenlinnantie 2  
FIN-00250 Helsinki, Finland  
E-mail: fredrik.almqvist@huch.fi

K. Ikäheimo · J. Piha  
Department of Child Psychiatry  
University of Turku  
Turku, Finland

K. Kumpulainen · I. Henttonen  
E. Räsänen  
Department of Child Psychiatry  
University of Kuopio  
Kuopio, Finland

S.-L. Linna · I. Moilanen  
Department of Child Psychiatry  
University of Oulu  
Oulu, Finland

K. Puura · T. Tamminen  
Department of Child Psychiatry  
University of Tampere  
Tampere, Finland

## Behavioural and emotional symptoms in 8–9-year-old children

**Abstract** We present epidemiological data from a multi-centre study on psychiatric symptoms among 6017 8–9-year-old children representing a total annual birth cohort (N = 60007) in Finland. The results are based on three questionnaires: the Rutter Parent Scale (RA2), the Rutter Teacher Scale (RB2), the Children's Depression Inventory (CDI). The proportion of children that scored above the cutoff points, indicating probable psychiatric disturbance, were 11.2% for the RA2, 13.9%

for the RB2 and 6.9% for the CDI. Twenty-four percent of the subjects scored above the cutoff point on at least one of the questionnaires. Low family social status and disrupted family relations correlated strongly with high rates of symptoms in the children.

**Key words** Childhood psychopathology – epidemiology – emotional symptoms – conduct symptoms – hyperactive symptoms – depressive symptoms

---

### Introduction

Information on the spectrum and prevalence of psychiatric symptoms and disorders, and their antecedents, is essential in order to promote psychiatric services for children, and to develop preventive measures in the field of child mental health (10). Psychopathology of a significant degree of severity and associated with impairment of functioning has been identified in 6–19%

of children and adolescents in studies from different parts of the world (1, 6, 8, 13, 29, 33, 34, 39, 47). According to a review by Bird (7), epidemiological studies earlier than 1980 reported a prevalence of childhood psychopathology in the range of 7–37%. In a recent comprehensive review by Verhulst (49), the reported prevalence of psychiatric dysfunction varied from 3–39%, with a median prevalence of 12%. If the definition of a disorder includes subjects with a mild level of disturbance, the overall rates

go toward 50%. Boys are reported to have higher rates of overall disturbance as well as disturbance of conduct, attention and hyperactivity, while emotional disorders are reported to be more frequent among girls. The findings are, however, strongly related to cultural and demographic factors and depend on the methods used in the studies.

Given the reasons mentioned above there is a need for well-founded national baseline information on the frequencies and spectra of psychiatric symptoms and disturbances in the child population. In order to be able to advocate and plan child psychiatric services, we also need reliable methods for screening, rating and assessing child psychiatric disorders, and information based on epidemiological research and scientific evaluation. To that end, in co-operation with all the five child psychiatric university units in Finland, we conducted a multicentre epidemiological study of mental health and disturbances in childhood. In the first part of the two-stage study the subjects were rated using three different screening questionnaires filled in by the child, his/her parents and teacher (2). In the second stage a sample of the children and their parents were interviewed (2, 3).

### Aims

The aims of this report were to estimate the rated frequencies of behavioural/emotional symptoms and to describe the spectrum of psychiatric symptoms in a child population also in relation to gender, social and familial factors. The interview-based estimate of the true prevalence of child psychiatric disorders will be published in another article of this supplement (3).

---

### Material

Finland has a population of 5 million inhabitants, and 93% speak with Finnish as their mother tongue. The target population of this study was school-age children represented by one annual age cohort. The basic population was the 60007 Finnish-speaking children born in Finland in 1981. A random sample of 6017 was drawn using a multiphase sampling procedure. This randomly picked sample constituted 10% of the whole age cohort, and the 5813 finally participating constituted 96.6% of the randomly selected sample.

Of the subjects, 50.6% were boys and 49.4% girls. Eighty-nine percent lived in a two-parent family, 83.4% lived with both biological parents. In the school year of autumn 1989 to spring 1990, 97.6% of the 8–9-year olds

attended primary school, 3.1% in the first grade, 96.6% in the second grade, 0.2% in the third grade, and 2.4% in special classes. Eighteen percent of the children came from the highest socio-economic classes, 62% from the two middle classes, and 20% from the two lowest classes. These figures are the same as the official census statistics (41–44).

A detailed description of the sampling procedure and the subjects, who represented the basic population extremely well, is published in another article of this supplement (2).

---

### Methods

We used the following scales: (1) the RA2 for completion by parents (32, 33) in a Finnish translation based on the English version A2, (4) the RB2 (31–35) for completion by teachers in a Finnish translation based on the English version B2. (5) These scales have been evaluated as valid in other cultures (13, 25, 26, 49) and have recently been reviewed (12). The items are summed up on each scale into a total score that measures overall deviance. Total scores of 13 or more on the RA2 and 9 or more on the RB2 indicate probable psychiatric disorder. Three subscales yield scores indicative of emotional or neurotic type disorders (5 items on the RA2, 4 items on the RB2), conduct or antisocial type disorders (5 items on the RA2, 6 items on the RB2) and hyperactivity (3 items on the RA2, 3 items on the RB2). A set of questions on demographic and family issues was given to the parents in connection with the RA2. (5) The CDI (14, 19, 20), shown to tap depressive symptoms in children (23, 28), was chosen as the children's self-report questionnaire. The scale was reviewed by Costello and Angold (9) and could be regarded as valid for our purposes, a view also supported by results published by Hodges (16). A comprehensive presentation of the inventory was published in 1992 by Kovacs (21). Our Finnish translation consists of 26 items. The question on suicidal ideation was excluded for ethical reasons, because the item could detrimentally upset the children.

After translation and re-translation the instruments were tested, modified and evaluated, and the procedures modified in a pilot study of 1000 8-year-old children (46). We recently published an evaluation of the three instruments, based on a ROC analysis, showing that the RA2 was a reliable screening instrument, and the RB2 an excellent one, and that the CDI, in combination with the

other two instruments, was a useful instrument for tapping overall psychiatric disturbance (22).

### Procedure

The parents' questionnaire was sent home from school with the child, and the child brought it back in a sealed envelope. If the questionnaire was not returned, a new one was sent to the parents twice, with a two-week interval. The teacher who knew the child best filled out the teacher's form. The child filled out the CDI in the classroom.

### Response

Only 204 children (3.4%) of the originally drawn sample of 6017 children did not participate in the study, due to refusal or failure to locate them. A small proportion of the questionnaires was never returned. Missing information on an item was replaced by the mean value of a) the mean of that item among the peer subjects in the same class and b) the mean of that item among the subjects in the same region, provided that the proportion of questions with missing information was less than one third of the total number of items on the scale. When there was no answer for one third or more of the questions, the questionnaire was excluded. Less than 1% of the returned forms had to be excluded. Thus the response rate was very high for all three questionnaires: 94.1% (n = 5664) for the RA2; 95.5% (n = 5746) for the RB2; and 94.5% (n = 5685) for the CDI.

### Statistical methods

The data was processed using the SAS software system for data analysis and applying the GLM and MIXED procedures, which accounts for variation and sampling weights in the five regional samples. The differences in distributions of the contingency tables were tested by means of the Chi-Square, or, when indicated, with other tests provided by the SAS.

The ethical committees at each of the five universities participating in the study approved the study.

## Results

### The RA2

The total score varied from 0–38 of a possible 62 for the 31 items. A proportion of 11.0% (that is closest to the

lowest 10<sup>th</sup> percentile) of the children scored 0 or 1, and 11.2% (closest to the highest 10<sup>th</sup> percentile) scored 13 or more points, the standard cutoff point for identifying

**Table 1** Mean total scores on the RA2 and the RB2

|                        | RA2<br>Mean (SE) | p <   | RB2<br>Mean (SE) | p <   |
|------------------------|------------------|-------|------------------|-------|
| Total score            |                  |       |                  |       |
| Boys                   | 7.16 (0.10)      |       | 5.07 (0.12)      |       |
| Girls                  | 5.88 (0.10)      | 0.001 | 2.38 (0.12)      | 0.001 |
| All                    | 6.52 (0.08)      |       | 3.72 (0.11)      |       |
| Emotional subscore     |                  |       |                  |       |
| Boys                   | 1.63 (0.04)      |       | 0.66 (0.03)      |       |
| Girls                  | 1.63 (0.04)      | ns    | 0.54 (0.03)      | 0.1   |
| All                    | 1.63 (0.04)      |       | 0.60 (0.03)      |       |
| Conduct subscore       |                  |       |                  |       |
| Boys                   | 1.14 (0.02)      |       | 1.46 (0.03)      |       |
| Girls                  | 0.63 (0.02)      | 0.001 | 0.33 (0.03)      | 0.001 |
| All                    | 0.89 (0.02)      |       | 0.89 (0.02)      |       |
| Hyperactivity subscore |                  |       |                  |       |
| Boys                   | 0.67 (0.03)      |       | 1.29 (0.03)      |       |
| Girls                  | 0.36 (0.03)      | 0.001 | 0.40 (0.03)      | 0.001 |
| All                    | 0.51 (0.02)      |       | 0.84 (0.02)      |       |

Data are based on 2862 boys and 2802 girls for RA2 and on 2910 boys and 2836 girls for RB2

**Table 2** The rated frequencies (%) of probable overall disturbance according to the RA2 and the RB2, and for different subtypes of disorders

|                 | RA2   |        | RB2   |        |
|-----------------|-------|--------|-------|--------|
|                 | %     | p      | %     | p      |
| Total*          |       |        |       |        |
| Boys            | 14.29 |        | 21.51 |        |
| Girls           | 7.99  | 0.001  | 6.17  | <0.001 |
| All             | 11.18 |        | 13.94 |        |
| Emotional **    |       |        |       |        |
| Boys            | 6.5   |        | 4.08  |        |
| Girls           | 4.5   | <0.001 | 2.47  | <0.001 |
| Total           | 5.5   |        | 3.29  |        |
| Conduct***      |       |        |       |        |
| Boys            | 5.6   |        | 15.63 |        |
| Girls           | 2.1   | <0.001 | 3.10  | <0.001 |
| Total           | 3.9   |        | 9.44  |        |
| Hyperactive**** |       |        |       |        |
| Boys            | 5.4   |        | 15.01 |        |
| Girls           | 2.1   | <0.001 | 2.82  | <0.001 |
| Total           | 3.8   |        | 9.00  |        |

\* Proportion scoring the cutoff point of 13 or more on RA2 or 9 or more on RB2.

\*\* Proportion scoring 13 or more on RA2 or 9 or more on RB2 and higher points on neurotic items than on antisocial items.

\*\*\* Proportion scoring 13 or more on RA2 or 9 or more on RB2 and higher points on antisocial than on neurotic items.

\*\*\*\* Proportion scoring 13 or more on RA2 or 9 or more on RB2 and 3 or more on the hyperactivity items.

Data are based on 2862 boys and 2802 girls for RA2 and on 2910 boys and 2836 girls for RB2.

**Table 3** Mean scores for depressive symptoms on the Children's Depression Inventory (CDI) and percentages for probable depressive disturbance when using 13 points and 17 points as cutoff points

|       | Mean (SE)   | cutoff= $\geq$ 13 (%) | cutoff= $\geq$ 17 (%) |
|-------|-------------|-----------------------|-----------------------|
| Boys  | 7.34 (0.17) | 16.88                 | 8.06                  |
| Girls | 6.37 (0.17) | 12.55                 | 5.31                  |
| All   | 6.86 (0.14) | 14.74                 | 6.70                  |
|       | $p < 0.01$  | $p < 0.001$           | $p < 0.001$           |

Data are based on 2880 boys and 2805 girls.

disturbance. The modal value was 4 ( $N = 523$ ), but 5 was very close ( $N = 519$ ). The median was between 5 and 6 points; 49.9% scored 0–5.9, and the rest scored 6 or more. The mean score was 6.52. The means for the partial emotional (4 items), conduct (5 items), and hyperactivity (3) scores are reported in detail in Table 1.

The mean of the total score on the RA2 was significantly greater for boys than for girls ( $p < 0.001$ ).

The gender difference was particularly high for the subscores of conduct disorder and hyperactive behaviour.

The frequency of probable disturbance was 11.2%, estimated by using the recommended cutoff point of 13 or more for disturbance, and it was significantly higher for boys than for girls. Further, the rates for conduct, emotional and hyperactive subtype disorders were higher in boys (Table 2).

#### The RB2

The total score varied from 0–37 of a possible 52 points on the 26 items. One third (32.2%) (closest to the lowest 10<sup>th</sup> percentile) scored 0 while 9.8% (closest to the highest 10<sup>th</sup> percentile) scored 10 or more. The modal value was 0. The median was 2, with 46.3% scoring 0–1.9 points, and the rest 2 or more. The mean was 3.72. The means for the partial scores are delineated in Table 1. The mean of the total score on the RB2 was significantly higher for boys than for girls. The boys also scored higher on all three subscores.

**Table 4** The frequencies (%) for probable disturbance according to the three questionnaires RA2, RB2 and CDI (scoring 17 or more) by social class and family type

|                     | RA2   |        |      | RB2   |        |      | CDI  |             |
|---------------------|-------|--------|------|-------|--------|------|------|-------------|
|                     | %     | (N)    | p    | %     | (N)    | p    | %    | (N)         |
| <i>Social class</i> |       |        |      |       |        |      |      |             |
| I (high)            | 7.0   | (996)  | ***  | 9.8   | (984)  | **   | 4.9  | (974) n.s.  |
| II                  | 9.7   | (1584) | n.s. | 11.5  | (1576) | n.s. | 5.6  | (1561) n.s. |
| III                 | 12.7  | (1880) | n.s. | 14.4  | (1868) | n.s. | 7.8  | (1849) n.s. |
| IV (low)            | 13.8  | (985)  | *    | 19.0  | (974)  | ***  | 6.9  | (966) n.s.  |
| V (mixed)           | 19.2  | (120)  | *    | 16.0  | (119)  | n.s. | 9.6  | (115) n.s.  |
| All                 | 11.2  | (5565) |      | 13.6  | (5521) |      | 6.6  | (5465)      |
|                     | $p <$ | 0.001  |      | 0.001 |        |      | 0.01 |             |
| <i>Family type</i>  |       |        |      |       |        |      |      |             |
| BBP                 | 9.3   | (4689) | ***  | 11.7  | (4651) | ***  | 6.2  | (4602) n.s. |
| BM & SF             | 23.3  | (271)  | ***  | 24.8  | (270)  | ***  | 9.3  | (269) n.s.  |
| BM                  | 18.1  | (519)  | ***  | 23.0  | (514)  | ***  | 8.1  | (508) n.s.  |
| BF & SM             | 24.0  | (25)   | n.s. | 16.0  | (25)   | n.s. | 4.0  | (25) n.s.   |
| BF                  | 11.3  | (53)   | n.s. | 32.1  | (53)   | ***  | 11.5 | (52) n.s.   |
| AP                  | 31.3  | (16)   | n.s. | 31.3  | (16)   | n.s. | 0.0  | (16) n.s.   |
| FP                  | 26.1  | (23)   | n.s. | 26.1  | (23)   | n.s. | 21.7 | (23) n.s.   |
| Other               | 36.4  | (22)   | **   | 31.8  | (22)   | n.s. | 13.6 | (22) n.s.   |
| All                 | 11.1  | (5618) |      | 13.7  | (5574) |      | 6.6  | (5517)      |
|                     | $p <$ | 0.001  |      | 0.001 |        |      | 0.01 |             |

\*\*\* =  $p < 0.0001$ ; \*\* =  $p < 0.001$ ; \* =  $p < 0.01$  – Mixed social class indicates a weak socioeconomic situation

BBP = The child living with both biological parents  
 BM & SF = The child living with biological mother and stepfather  
 BM = The child living with biological mother only  
 BF & SM = The child living with biological father and stepmother  
 BF = The child living with biological father only  
 AP = The child living with adoptive parents  
 FP = The child living with foster parents  
 Other = The child living with others, most often in children's home

The frequency of probable disturbance on the RB2 was 13.9%; it was estimated by using the recommended cutoff point of 9 or more, and was over three times higher for boys than for girls. The boys were also rated higher on frequencies of emotional, conduct and hyperactive subtypes of disorder (Table 2).

#### Self-report of depressive symptoms (CDI)

The total score varied from 0–48 of a possible 52 points for the 26 items. A proportion of 14.5% (closest to the lowest 10<sup>th</sup> percentile) scored 0–1 and 10.3% (closest to the highest 10<sup>th</sup> percentile) scored 15 or more. The modal value was 3. The median was 6 and 50% scored 0–5.9, the rest 6 or more. The mean score was 6.7, and significantly higher for boys than for girls (Table 3).

The frequency of self-reported depression was 6.7% when we used the cutoff point of 17 points or more, which is one of the higher alternative cutoff points recommended in the literature (21). The rate was significantly higher for boys (8.1%) than for girls (5.3%) (Table 3). A lower recommended cutoff point (13 points) yielded a frequency of 14.7%, and it also revealed a significant difference between boys (16.9%) and girls (12.6%).

#### Background variables

Significant differences in disturbances were noted in children belonging to different socio-economic classes and for children living in families with different structures. Going from the most favourable social class, I, to the most disadvantaged class, IV, (and the mixed class V), the percentage of disturbed children showed a very significant rise on all the three screening instruments.

Children living in families with both biological parents had significantly lower frequencies of disturbance than almost any other family constellation on all the three scales (Table 4). The findings were the same for the emotional, conduct and hyperactivity subtypes of disorder on both Rutter's scales. The weakest trend was noted for emotional disorder on the RA2; an increase from 4.1% in social class I to 6.8% in class IV ( $p < 0.05$ ). According to the CDI children who lived with their biological father and a stepmother, and adopted children, had lower frequencies of depression than children living with both biological parents.

The frequency of probable disturbance according to the RA2 was significantly ( $p < 0.01$ ) lower in stable housing areas (9.2%) than in growing areas with many people moving in (11.7%), and in declining areas with

**Table 5** Pairwise comparison of the proportions (total %) of overlapping between screen + and screen –

| RA2 score  | RB2 score              |           |         | N    |
|------------|------------------------|-----------|---------|------|
|            | <9 (%)                 | =/≥9 (%)  | Sum (%) |      |
| <13        | 79.3                   | 9.7       | 89.0    | 5603 |
| =/≥13      | 7.1                    | 3.9       | 11.0    |      |
|            | 86.4                   | 13.6      | 100.0   |      |
|            | Total score on the CDI |           |         |      |
| RA2; score | <17 (%)                | =/≥17 (%) | Sum (%) |      |
| <13        | 84.3                   | 4.8       | 89.1    | 5545 |
| =/≥13      | 9.1                    | 1.8       | 10.9    |      |
|            | 93.4                   | 6.6       | 100.0   |      |
|            | Total score on the CDI |           |         |      |
| RB2 score  | <17 (%)                | =/≥17 (%) | Sum (%) |      |
| <9         | 81.8                   | 4.5       | 86.3    | 5678 |
| =/≥9       | 11.5                   | 2.2       | 13.7    |      |
|            | 93.3                   | 6.7       | 100.0   |      |

many people moving out (12.9%). A slightly higher ( $p < 0.05$ ) frequency was observed in urban town areas (12.0%) when compared with suburban (10.1%) and rural districts (10.7%). No significant difference was observed between the five regional study areas ( $p = 0.2$ ).

According to the RB2 the frequency of probable disturbance was lower ( $p < 0.01$ ) in suburban areas (11.0%) than in urban (14.6%) and rural districts (14.7%), but there were no significant differences between the various types of housing areas or between the five regional study areas.

According to the CDI the frequency of probable depression was slightly ( $p < 0.05$ ) higher in growing housing areas (7.2%) than in stable (5.9%) and declining (5.9%) areas, but there were no significant differences between urban, suburban and rural districts or between the five regional study areas.

#### Overlapping between the screens

Pairwise comparisons between the screening measures (Table 5) revealed that 3.9% of the children were identified as probably disturbed by both the RA2 and RB2, and an additional 16.8% by either of the instruments, a total of 20.7% for both scales in combination.

Of the children who scored above the cutoff point on the RA2, 36% also scored above the cutoff point on the RB2, and of those who scored above the cutoff

**Table 6** Cross-tabulation of the ratings on the RA2 and RB2 and on the CDI for the 5541 children with accepted form for all three instruments

| RB2<br>Total score |        | Total score on the CDI |        |      |
|--------------------|--------|------------------------|--------|------|
|                    |        | <17                    | =/>>17 | Sum  |
| <9                 | (tot%) | 85.1                   | 4.2    | 89.3 |
|                    | (row%) | 95.4                   | 4.7    |      |
|                    | (col%) | 90.0                   | 77.1   |      |
| =/>9               | (tot%) | 9.5                    | 1.2    | 10.7 |
|                    | (row%) | 88.5                   | 11.5   |      |
|                    | (col%) | 10.0                   | 22.9   |      |
| Total              | 94.6   | 5.4                    | 100.0  |      |

Chi-Square: Value = 43.7, DF = 1,  $p < 0.001$ Children ( $N = 603$ ) with total score on RA2 =/>>13

| RB2<br>Total score |        | Total score on the CDI |        |      |
|--------------------|--------|------------------------|--------|------|
|                    |        | <17                    | =/>>17 | Sum  |
| <9                 | (tot%) | 57.6                   | 6.8    | 64.3 |
|                    | (row%) | 89.4                   | 10.6   |      |
|                    | (col%) | 68.8                   | 41.4   |      |
| =/>9               | (tot%) | 26.0                   | 9.6    | 35.7 |
|                    | (row%) | 73.0                   | 27.0   |      |
|                    | (col%) | 31.2                   | 58.6   |      |
| Total              | 83.6   | 16.4                   | 100.0  |      |

Chi-Square: Value = 27.2, DF = 1,  $p < 0.001$ 

point on the RB2, 29% also scored above the cutoff point on the RA2. Only 16% of the children who scored above the cutoff point on the RB2 also scored above the cutoff point on the CDI, and only 17% of the children who scored above the cutoff point on the RA2 also scored above the cutoff point on the CDI. Of the children scoring 17 or more on the CDI, 33% were rated as disturbed on the RB2, and 27% rated as disturbed on the RA2.

Crosstabulation of all three screens are presented in Table 6 in detail. Of the 5541 children screened by all three instruments, 4203 (76%) were rated as probably healthy by all three, and 1338 children (24%) were rated as probably disturbed by at least one scale. Only 1% of the children was rated in all three questionnaires as probably disturbed (Table 7).

The use of multiple methods allowed us to test how the children from whom we did not receive any response on one screen were classified according to the ratings on the two other screens. We did not receive any response on the RA2 from 80 (1.4%) of the 5746 children that responded to the RB2, and in 110 cases (1.9%) we received the

**Table 7** Overlapping between the three screens RA2, RB2 and CDI for 5541 children with adequate responses to all three screens. The figures are total % distribution completing 100%

|                 |                  | Depression by the CDI |          |
|-----------------|------------------|-----------------------|----------|
|                 |                  | No                    | Yes      |
| Parent's<br>RA2 | Teacher's<br>RB2 | n                     | n        |
| Healthy         | Healthy          | 76% (4203)            | 4% (205) |
| Healthy         | Disturbed        | 9% (469)              | 1% (61)  |
| Disturbed       | Healthy          | 6% (347)              | 1% (41)  |
| Disturbed       | Disturbed        | 3% (157)              | 1% (58)  |

response only after a reminder. The rate of disturbance according to the RB2 was 30% for the RA2 "non-responders", and 23.6% for those who responded after the reminder, compared with 13.5% for those who responded in the first round ( $p < 0.001$ ). The same phenomenon was also observed for "nonresponders" and "reminders" on the RB2 and the CDI according to their ratings by the two other screens. Of the 5731 children that responded to the RB2, the rate of disturbance was 26.7% among the 15 children who did not respond to the CDI compared with 13.8% among those who responded at once ( $p < 0.001$ ).

## Discussion

### Material

The subjects made up a large, representative sample of all children born in Finland during one year. The overall dropout rate due to migration, refusal and incorrect response was small (5.9%), and did not influence the validity of the findings. The subjects with missing information on one screen were, in the light of the ratings on the two other screens, evidently more disturbed than the respondents. Due to the small proportion of nonrespondents (0.04% for the RB, 0.2% for the CDI, and 1.4% for the RA) this did not, however, significantly reduce the reported rates. It is also important to note that the rate of disturbances among those who responded only after reminders was much greater than among those that responded directly. These findings have implications on the interpretation of the results in any study with even moderate dropout rates.

The demographic characteristics of the subjects corresponded closely with census statistics, which means that the subjects were a good representation of the total pop-

ulation of 8–9-year-old children and their families in Finland. As a consequence of the slightly different sampling ratios and demographic structure in the five regions, the subjects in the study do not, however, exactly represent the age cohort of the entire country. Hence the results presented here are weighted and representative for the whole country.

## Methods

The proportion of refusal was 1.1%. Missing information on items of the returned questionnaires was generally below 0.5%, and was at most 1–2% for a few items. No major confusion or difficulties in understanding were observed among the children when answering the depression inventory at school. A small number of children needed assistance from the teacher or researcher.

The procedure of substituting missing values with the mean of class and regional peers in questionnaires with less than one third nonresponded items on the standardized scales had to be applied in only 8% of the RA2, 3% on the RB2 and 5% on the CDI. Thus, in this sense, the accuracy of the information from the questionnaires has to be regarded as very high.

## Rating scales

Parents and teachers are the first to recognize emotional and behavioural disturbances in children. Referral for psychiatric evaluation, consultation or treatment is usually initiated by both of them. Parents and teachers are important informants, because they provide different information on the children as they experience them in different ways, and see them in diverse situations (11, 15, 33, 38). For these reasons the engagement of both parents and teachers as informants is well motivated.

The use of the child's self-report questionnaire (CDI) turned out to be justified, as it expanded the spectrum of symptoms, and increased the total proportion of children rated as probably disturbed. In contrast to the RA2 and RB2, which are global measures for psychopathological symptoms, the CDI is a method for tapping depressive symptoms. The 27 items of the CDI have, however, been shown to constitute five factors: negative mood, interpersonal problems, ineffectiveness, anhedonia, and negative self-esteem (21). Accordingly, the CDI also measures a broad range of psychopathological symptoms.

It has to be emphasised that the children who scored above the cutoff points are “probably disturbed”. The “true” rates for disturbances will be reported in a following paper where we evaluate the scales on the basis of structured diagnostic interviews with the parents (3).

## Results

### *The RA2*

The frequencies of children who scored 13 or above on the RA2 in our study were the same as those reported earlier in normal unselected populations, i.e., about 15% for boys and 8% for girls in England (31, 33). The frequencies among 12–15-year-old children in Japan were 11% for boys and 9.4% for girls (27). In a collaborative study in the Western Pacific regions the rate for both sexes was 7% in China, 12% in Japan and 19% in Korea (25, 50). A high rate has been reported from Greece (39%) (30).

Contrary to earlier studies the proportion of emotional disturbances in our study was higher among boys than among girls. Emotional disturbance was only slightly higher than conduct disturbance in boys, but emotional disturbance in girls was more than two times more frequent than conduct disturbance. The rate of hyperactivity (3.8%) was much lower than that (9.9%) reported by Schachar et al. (38), but the male predominance is in line with earlier reported data.

By using the earlier reported figures of sensitivity (55%) and specificity (94%) (4, 12, 33), the calculated true frequency of psychiatric disturbance reached 18%.

### *The RB2*

The proportions of children rated by the RB2 as probably disturbed were of the same size as in earlier reports. The earliest studies reported somewhat lower rates, in Aberdeen 9–11% for boys and 3–5% for girls (31), and in the first Isle of Wight study 10% for boys and 4% for girls (33). The rates were also lower in the study from Japan, 10.4% for boys and only 4.0% for the girls (27). In the West Pacific collaborative study the rate was 4% in Japan, 8% in China and 14% in Korea (25, 50).

Higher frequencies have been reported, especially in later studies. In Sweden, 17% of 11–13-year-old children were rated as disturbed (45). In a second comparative study of children on the Isle of Wight and in London, 24.5% of the boys and 13.2% of the girls in London, and 13.8% of the boys and 7.1% of the girls on the Isle of Wight were rated as disturbed (34). Higher frequencies of at least 20–30% have also been reported (13, 18, 30, 48).

The proportions of both emotional and conduct disturbances were higher in boys than in girls. The conduct

features were clearly greater than the emotional features in boys, but only slightly in girls, a result that corresponds well to earlier findings. In the London study of 10-year-old children a greater proportion of the boys (8%) than of the girls (7%) were designated as emotionally disturbed. The proportion of children with hyperactive disorders (9.0%) was close to that (8.2%) reported by Schachar et al. (38).

In our study, the earlier reported figures of sensitivity (53%) and specificity (93%) (5, 12, 33) revealed a calculated true frequency for psychiatric disturbance of 17%.

### *The CDI*

Our total mean score was 6.7 and is slightly higher than the mean score of 6.1 reported by Kronenberg et al. (23) for healthy children, and Kaslow et al. (17) for normal school children. It is at the same level as that reported by Saylor et al. (37), slightly higher compared with children of the same age in Sweden (24), but lower than that (9.8) reported by Kovacs (21), and that (8.6) reported by Smucker et al. (40). Like Kovacs (21), we found a significantly higher mean for boys than for girls. On the contrary, Larson and Melin (24) reported a significantly higher mean for girls, and Smucker et al. (40) reported no sex differences in this age group.

The exclusion of item 9, "thinking of killing myself", in this study does not significantly interfere with a comparison of the results of this study and earlier published data on the CDI, because of the low rate of positive answers for this item in an unselected population sample.

The cutoff point of 15 was closest to the upper 10<sup>th</sup> percentile (in our study 10.3%). For children of a similar age, Larsson and Melin (24) reported the upper 10<sup>th</sup> percentile cutoff at 13, Kovacs (10) reported the upper 10<sup>th</sup> percentile cut off at 20 points, and Smucker et al. (40) at 19.

There is no generally accepted cutoff point for tapping depressive disorders in epidemiological research. In a clinical study, 15 was the optimal cutoff score for discrimination between depressed and non-depressed children (12). The figures for sensitivity (54%) and specificity (84%) at the cutoff score 19, reported by Hodges (16), give a calculated "true" prevalence of 7% for depression in this study. It needs to be noted, however, that the CDI has been reported to tap a broader spectrum of psychopathology (especially anxiety) than depression alone (36).

Our results, which show that boys have higher rates of internalising symptoms, are partly contrary to earlier

reported findings. Representative population samples, however, have shown that prevalences of internalising and depressive symptoms in preadolescent and especially younger school age boys are higher than in girls. Later in adolescence and young adulthood the internalising symptoms are clearly higher in females. In clinical settings externalising symptoms dominate among boys in all age groups. One explanation might be that parents and teachers, and even clinicians, do not pay proper attention to symptoms of anxiety and depression in boys. These symptoms might either spontaneously heal, or they might, if not cured, gradually turn into behaviour disturbances and later emerge as well known externalising symptoms.

For the RA2, RB2 and CDI, there were very significant variations in rated frequencies according to social class and family constellation, along with the gender differences. This may partly explain the differences between earlier studies in frequencies of disturbances among child populations with various cultural, demographic and social compositions. In our study, differences were also noted to correlate with the degree of urbanisation, and the stability of the housing area. But when we compared the five regional study areas, there were no significant differences. These findings indicate that the correlation between psychopathology and socio-demographic factors is more likely to relate to family factors than to broader social and demographic factors, at least in a culturally homogenous country like Finland where there is only one city (Helsinki) with a population of almost half a million.

### *Overlapping*

It is to be expected that only some of the children rated as disturbed on one scale were also rated as disturbed on another scale. This is partly due to the fact that the scales, particularly the CDI in relation to the RA2 and the RB2, are designed to tap different aspects of childhood psychopathology. However, even when the focus of two of the scales are in the same area of the spectrum of disturbances, as is the RA2 and the RB2, it is to be expected that the scales will not overlap completely. The parent has a different view of the child, a different emotional relationship to the child. The parents see the child in different situations and have other expectations of the child than the teacher. The child also behaves differently in different situations and contexts.

The 4% overlap between the RA2 and RB2 in our study can be compared to corresponding figures reported from a study in Japan: 2.3% (27) and another from the same region: Japan; 1.4%, China; 2.1% and Korea; 4.5% (25). Overlapping between the CDI, which taps depres-



sion (and is more specific than the globally oriented Rutter scales), and the Rutter scales, has not been reported earlier.

The three rating scales were supplements to each other, contributing to the building up of a more comprehensive assessment of the psychiatric problems of the children, and together provided a more complete picture of the spectra of psychopathology in childhood. The proportion of children rated as probably disturbed by at least one of the three instruments was high, but not unrealistically high in view of the earlier mentioned moderate sensitivity, but rather high specificity, and the high cutoff point for the CDI used in this study.

It is surprising that only 1% of all the children were rated as probably disturbed on all three instruments. These children constitute 4.3% of those scoring high on any of the three instruments. Nevertheless, if a child was selected by one instrument the likelihood of he/she also being selected by another instrument was greatly increased. The majority of children rated as disturbed on one screen were, however, rated as healthy by the others.

If, as has been suggested (21, 24), the CDI is a valid method for screening depression and emotional disturbance in children, we would have missed a substantial proportion (205/1338), or 15.3% of the children rated as probably disturbed if we had not used the CDI in our study.

The teachers' scale was slightly more accurate than the parents' scale in finding the children with subjective depressive symptoms (33% versus 29%). It is, however, not correct to say that teachers are more aware than parents of children with depression, because with the traditional cutoff points a greater proportion of the children were rated as disturbed by the RB2 (14%) than by the RA2 (11%). If we used 12 as a cutoff point for the RA2, 14% of all children were rated as disturbed, and the RA2 identified 32% of the children scoring 17 or more on

the CDI. Thus, in this case, the Rutter scales are just as efficient in identifying depressed children.

## Conclusions

Each of the three screening instruments makes an important contribution to the picture of the mental disturbance of a single child, and to the epidemiology and spectra of psychiatric symptoms in childhood. The validity of the role of the single instruments still has to be evaluated by information based on interviews with the child and his parents.

The rated frequencies of dysfunctional levels of total psychiatric symptomatology are lower than the median (12%) reported by Verhulst (49), when based on one instrument, and higher when based on a combination of two scales. The rate of 24% resulting from the combination of three scales is rather high, but still below many of the earlier reported prevalence rates for psychiatric disturbance in population based studies.

The findings also indicate that adults do not always know enough about a child's feelings and state of mood. Children who do not worry adults with their symptoms may suffer without adults' recognising their problems.

On the basis of earlier published data on the validity, sensitivity and specificity of the three methods used, it is obvious that the true prevalence of psychiatric disturbance might be close to 20% in school-age children. Since these results are based on screening instruments, they have to be confirmed by results based on interviews. Further analysis of the information from the interview stage of the study will throw more light on the screened results and complement the picture of childhood psychopathology.

**Acknowledgement** The S. and A. Gyllenberg Foundation supported the study.

## References

1. Achenbach TM, Howell CT (1993) Are American children's problems getting worse? A 13-year comparison. *Journal of the American Academy of Child and Adolescent Psychiatry* 32:1145–1154
2. Almqvist F, Ikäheimo K, Kumpulainen K, Tuompo-Johansson E, Linna S-L, Puura K, Moilanen I, Räsänen E, Tamminen T, Piha J (1999) Design and subjects of a Finnish epidemiological study on psychiatric disorders in Childhood. *European Child and Adolescent Psychiatry* 8: Suppl. 1 (this issue)
3. Almqvist F, Puura K, Kumpulainen K, Tuompo-Johansson E, Henttonen I, Huikko E, Linna S-L, Ikäheimo, Aronen E, Katainen S, Piha J, K Moilanen I, Räsänen E, Tamminen T, (1999) Psychiatric disorders in 8–9 year-old children based on a diagnostic interview with the parents. *European Child & Adolescent Psychiatry, Supplement* (this issue)
4. Almqvist F, Tuompo-Johansson E, Panelius E, Aronen E, Kairemo A-C (1991) Screening for psychiatric symptoms on the basis of the Rutter Parents questionnaire. In: Tamminen T, Almqvist F, Piha J (Eds) *Psychiatric Symptoms in Finnish Children. Reports of Psychiatria Fennica* 91: 21–33
5. Almqvist F, Tuompo-Johansson E, Panelius E, Aronen E, Kairemo A-C (1991) Screening for psychiatric symptoms on the basis of the Rutter Teacher questionnaire. In: Tamminen T, Almqvist F, Piha J (Eds) *Psychiatric Symptoms in Finnish Children. Reports of Psychiatria Fennica* 91: 35–46

6. Anderson J, Williams S, McGee R, Silva P (1987) DSM-III disorders in preadolescent children. *Archives of General Psychiatry* 44:69–76
7. Bird HR (1996) Epidemiology of childhood disorders in a cross-cultural context. *Journal of Child Psychology and Psychiatry* 37:35–49
8. Bird HR, Gould MS, Yager T, Staghezza B, Canino G (1989) Risk factors for maladjustment in Puerto Rican children. *Journal of the American Academy of Child and Adolescent Psychiatry* 28:847–850
9. Costello E, Angold A (1988) Scales to assess child and adolescent depression. Checklists, screens and nets. *Journal of the American Academy of Child and Adolescent Psychiatry* 27:726–737
10. Costello EJ, Burns BJ, Angold A, Leaf PJ (1993) How can epidemiology improve mental health services for children and adolescents? *Journal of the American Academy of Child and Adolescent Psychiatry* 32:1106–1113
11. Cox A, Rutter M (1985) Diagnostic appraisal and interviewing. In: Rutter M, Hersov L (Eds) *Child and Adolescent Psychiatry. Modern Approaches* (pp. 233–250). London: Blackwell Scientific Publications
12. Elander J, Rutter M (1996) An update on the status of the Rutter Parents' and Teachers' scales. *Child Psychology & Psychiatry Review* 1:31–35
13. Fombonne E (1994) The Chartres study: I. Prevalence of psychiatric disorders among French school-aged children. *British Journal of Psychiatry* 164:69–79
14. Fundudis T, Berney T, Kolvin I, Famuyiwa O, Barrett, L, Bathe S, Tyrer S (1991) Reliability and validity of two self-rating scales in the assessment of childhood depression. *British Journal of Psychiatry* 159:Suppl 11, 36–40
15. Graham P (1967) Perceiving disturbed children. *Special Education* 56:29–33
16. Hodges K (1990) Depression and anxiety in children: A comparison of self-report questionnaires to clinical interview. *Psychological assessment. American Journal of Consulting and Clinical Psychology* 2:376–381
17. Kaslow N, Rehm L, Siegel A (1984) Social-cognitive correlates of depression in children. *Journal of Abnormal Child Psychology* 12: 605–620
18. Kolvin I, Garside R, Nicol R, Leitch I, Macmillan A (1977) Screening schoolchildren for high risk of emotional and educational disorder. *British Journal of Psychiatry* 131: 192–206
19. Kovacs M (1981) Rating scales to assess depression in school-aged children. *Acta Paedopsychiatrica* 46: 305–315
20. Kovacs M (1985) The childrens depression inventory. *Psychopharmacological Bulletin* 21:995–998
21. Kovacs M (1992) *Childrens Depression Inventory, CDI, Manual*. Ontario: Multi-Health Systems Inc.
22. Kresanov K, Tuominen J, Piha J, Almqvist F (1998) Validity of child psychiatric screening methods. *European Child and Adolescent Psychiatry* 7:85–95
23. Kronenberg Y, Blumensohn R, Apter A (1988) A comparison of different diagnostic tools for childhood depression. *Acta Psychiatrica Scandinavica* 77:194–198
24. Larsson B, Melin L (1992) Prevalence and short-term stability of depressive symptoms in schoolchildren. *Acta Psychiatrica Scandinavica* 85:17–22
25. Matsuura M, Okubo Y, Kojima T, Takahashi R, Wang Y, Shen Y, Lee C (1993) A cross-national prevalence study of children with emotional and behavioural problems – a WHO collaborative study in the Western Pasific region. *Journal of Child Psychology and Psychiatry* 34:307–315
26. Morita H, Suzuki M, Kamoshita S (1990) Screening measures for detecting psychiatric disorders in Japanese secondary school children. *Journal of Child Psychology and Psychiatry* 31:603–617
27. Morita H, Suzuki M, Suzuki S, Kamoshita S (1993) Psychiatric disorders in Japanese secondary school children. *Journal of Child Psychiatry and Psychology* 34:317–332
28. Nelson W, Politano P, Finch A, Wendel N, Mayhall C (1987) Children's depression inventory: Normative data and utility with emotionally disturbed children. *Journal of the American Academy of Child and Adolescent Psychiatry* 26:43–48
29. Offord DR, Boyle MH, Racine Y (1989) Ontario child health study. Correlates of disorder. *Journal of the American Academy of Child and Adolescent Psychiatry* 28:856–860
30. Papatheophilou R, Bada K, Micheliyamiannakis I, Makaronis G, Pantelakis S (1981) Psychiatric disorders in 6- to 8-year old children in the greater Athens area. *Bibliotheca Psychiatry* 160:92–100
31. Rutter M (1967) A children's behaviour questionnaire for completion by teachers: Preliminary findings. *Journal of Child Psychology and Psychiatry* 8:1–11
32. Rutter M, Graham P (1966) The reliability and validity of psychiatric assessment of the child: I. Interview with the child. *British Journal of Psychiatry* 114:563–579
33. Rutter M, Tizard J, Whitmore K (1970) (Eds) *Education, Health and Behaviour*. London: Longman
34. Rutter M, Cox A, Tupling C, Berger M, Yule W (1975) Attainment and adjustment in two geographical areas. I. The prevalence of psychiatric disorder. *British Journal of Psychiatry* 132:6: 493–509
35. Rutter M, Tizard J, Yule W, Graham P, Whitmore K (1976) Isle of Wight studies, 1964–1974. *Psychological Medicine* 6:313–332
36. Saylor C, Finch A, Baskin C, Saylor B, Darnell G, Furey W (1984) Children's depression inventory: Investigation of procedures and correlates. *Journal of the American Academy of Child Psychiatry* 23:626–628
37. Saylor C, Finch A, Spirito A, Bennett B (1984) The childrens's depression inventory: A systematic evaluation of psychometric properties. *Journal of Consulting and Clinical Psychology* 52:955–967
38. Schachar R, Rutter M, Smith A (1981) The characteristics of situational and pervasive hyperactive children: Implications for syndrome definition. *Journal of Child Psychology and Psychiatry* 22:375–392
39. Schwartz Gould, M, Wunsch-Hizig R, Dohrenwend B (1981) Estimating the prevalence of childhood psychopathology. A critical review. *Journal of the American Academy of Child and Adolescent Psychiatry* 20:462–476
40. Smucker M, Craighead E, Craighead L, Green B (1986) Normative and reliability data for the Children's Depression Inventory. *Journal of Abnormal Child Psychology* 14:25–39
41. Statistics Finland, Report 10, 1988
42. Statistics Finland, Population, 7, 1992
43. Statistics Finland, Education, 3, 1992
44. Statistics Finland, Education, 3, 1993
45. Svedhem L (1991) *Socialt nätverk och beteendeproblem i skolan hos 11–13 åringar*. Stockholm: Carlsson Bokförlag
46. Tamminen T, Almqvist F, Piha J (Eds) (1991) *Psychiatric Symptoms in Finnish Children. Reports of Psychiatria Fennica No 91*
47. Velez CN, Johnson J, Cohen P (1989) A longitudinal analyses of selected risk-factors for childhood psychopathology. *Journal of the American Academy of Child and Adolescent Psychiatry* 28:861–864
48. Venables P, Fletcher R, Dalias J, Mitchell D, Schulsinger F, Mednick S (1983) Factor structure of the Rutter "children behaviour questionnaire" in a primary school population in a developing country. *Journal of Child Psychology and Psychiatry* 24:213–222
49. Verhulst, FC (1995) Review of community studies. In: Verhulst FC, Koot HM (Eds) *The Epidemiology of Child and Adolescent Psychopathology*. New York: Oxford University Press
50. Wong C (1988) The Rutter Parent Scale A2 and Teacher Scale B2 in China II. Clinical validity among Chinese children. *Acta Psychiatrica Scandinavica* 78:11–17