Peers, teachers and parents as assessors of the behavioural and emotional problems of twins and their adjustment: the Multidimensional Peer Nomination Inventory

Lea Pulkkinen1, Jaakko Kaprio2 and Richard J Rose3

1University of Jyväskylä, Finland
2University of Helsinki, Finland
3Indiana University, Bloomington, USA

A multidimensional peer nomination inventory (MPNI; 30 items) and parallel versions – MPNI-Teacher and MPNI-Parental Rating Forms (37 items each) – were developed during a major new Finnish study of families with twins. The twins (477 girls and 467 boys) were 12 years old, representing subsets of three nationwide Finnish twin cohorts (b. 1983–1985). They were enrolled in 503 school classes, and the total number of children participating in peer nominations was 12,937. Three main factors were extracted from peer nominations and teacher and parental assessments. Intercorrelating sub-components were found, especially in parental assessments. Scales were formed, accordingly, for Behavioral Problems (including Hyperactivity-Impulsivity, Aggression, and Inattention), Emotional Problems (including Depression and Social Anxiety), and Adjustment (including Constructiveness, Compliance, and Social Activity). A framework for the development of the multidimensional inventory was a model of emotional and behavioural regulation. Peer nominations were most reliable, while parental assessments, although mostly satisfactory, were least reliable. Results provided evidence of concurrent validity of peer-referenced assessment, using teacher assessments as criteria; correlations between assessments of peers and parents were lower. The inventory has discriminative validity. Intra-pair correlations of monozygotic co-twins were higher than correlations of same-sex (SS) and opposite-sex (OS) dizygotic (DZ) co-twins for all scales across all three assessors, and peer nominations of both SS and OSDZ co-twins yielded correlations significantly greater than zero for all scales. All scales, except Depression and Social Anxiety, differentiated boys from girls.

Keywords: twin study, behavioral disorders, disturbance, emotional, social adjustment, social behavior, assessment, personality

Introduction

The first aim of the present study was to develop a multidimensional peer nomination inventory, with parallel versions for teacher and parental assessments, for the study of children’s social behaviour, and to investigate concurrent and discriminative validity of these assessments with a new Finnish twin sample. As these are the first data reported from the new Finnish twin study, a second aim is to introduce that current study, which we call FinnTwin12.

Available peer nomination inventories are limited in content. They have been developed for assessment of aggression,1 prosocial behaviour,2 depression,3 or children’s peer status.4 Peer nomination content varies from study to study, and no peer nomination inventory which covers diverse dimensions of social behaviour, and which has been used in diverse studies, is available.

The peer nomination method is based on the Guess-Who technique originally presented by Hartshorne and May in 1929,5 in which children’s social reputations were assessed by their peers. Peer nomination with multiple informants is a technique that can detect characteristics of children’s social behaviour not evident to adults. With multiple peer informants, assessments are not sensitive to generalised conceptions that may bias assessments by adults.

Peer assessments have been obtained by two procedures. In one, all classmates are assessed on all studied characteristics contained in a given scale6 (eg ‘Starts bullying’ rated on a three-point scale: 1 = never, 2 = sometimes, 3 = often). In the alternate procedure, peers choose classmates who best fit a
given characteristic (e.g. ‘Which of your classmates bully smaller and weaker kids?’). The latter procedure, called peer nomination, is less demanding and less time consuming than peer assessment, since assessors concentrate on the ‘very typical’ end of the assessment scale. By calculating the ‘votes’ given by classmates to each individual (and corresponding percentages of the maximum number of votes), a ranking order of individuals can be created for each studied characteristic. The concurrent validity of peer nominations, using teacher assessments as criteria, has proved encouraging.

Most characteristics studied in peer nominations or peer assessments have concerned overt behaviour. Few peer nomination studies have focused on childhood anxiety and depression, although anxieties and fears may be quite prevalent in school-age children, and a need for the classification of childhood depression has been recognised. It has been questioned how reliably children can assess internalising behaviours, but some findings show that children are indeed able to identify symptoms of anxiety and depression in their peers.

The present study continues a methodological development started during the Jyväskylä Longitudinal Study of Personality and Social Development, conducted by Pulkkinen. The study began in 1968, when a multidimensional 33-item peer nomination inventory and its parallel teacher assessment version were developed for the study of 8-year-old children’s social behaviour. A framework for the selection of items was a two-dimensional impulse control model, recently modified and relabelled as a model of emotional and behavioural regulation (Figure 1). The development of the multidimensional peer nomination inventory (MPNI), and its teacher and parental rating forms was continued in the same framework, because empirical findings based on peer nominations and teacher assessments have supported the assumptions of the circumplex model. Data have confirmed the predictive value of the model over 20 years. Good self-control in childhood predicted good social adjustment at a later age. Conversely, children with problems in self-control were more likely than children with good self-control to experience accumulated problems in social functioning during the course of their development. Problems in social functioning included alcohol abuse, antisocial behaviour, and an unstable employment history.

Emotional regulation helps maintain internal arousal within a manageable, optimal performance range, whereas behaviour regulation helps adjust reactions to external circumstances. As described elsewhere, the model includes inhibitory and enhancing processes: the neutralisation and intensification of emotion and the suppression and activation of behaviour. The model suggests two major types of behaviour characterised by poor self-control, Type A and D, parallel to the most frequently referred dimensions of children’s social behaviour, externalising and internalising behaviour, respectively. Type C (eg compliance) means behaviour where both emotion and behaviour are controlled or inhibited. It differs from the other type of inhibited behaviour, Type D (eg anxiety), in the control of emotion. In Type C, emotional balance is maintained, whereas emotional imbalance or undercontrol is typical of Type D. Type B (eg constructiveness), on the other hand, includes overt behaviour with well-controlled emotion. An individual behaving constructively considers other people’s needs and rights, and adjusts their own needs and rights to offer a mutually acceptable outlet from an emotionally arousing situation. It differs from the other type of overt behaviour, Type A (eg aggression), in the control of emotion parallel to Compliance and Anxiety: in Type B, emotional balance is maintained, whereas emotional imbalance or low control is typical of Type A.

Pulkkinen’s original peer nomination inventory included items for different types of aggression (direct, indirect, proactive, reactive, physical, verbal, and facial), social anxiety, compliance, and constructiveness, and for the ‘reference axis’, poor

---

Figure 1 A model of emotional and behavioral regulation (Pulkkinen, 1995)
and weak self-control and social activity and passivity. These items form the core of the revised multidimensional inventory presented here, although some items were omitted, new items added, and some were modified in wording. Items for indirect aggression were changed to emphasize social manipulation ('Goes round telling people's secrets to others') instead of outbursts of anger ('Displays aggression toward some object'), because girls tend to avoid situations where direct confrontation is required to solve problems, for example, by manipulating the peer group.25–27

Additional items were added to include hyperactivity-impulsivity, inattention, and depression in the inventory. Based on a study by McGee and Williams,28 depression was expected to correlate with anxiety and represent TypeD behaviour of the model. Correspondingly, hyperactivity was expected to correlate with aggression as found in several studies29–31 and represent TypeA behaviour of the model. In the Diagnostic and Statistical Manual of the American Psychiatric Association (DSM IV, 1994), Attention-Deficit/Hyperactivity Disorder is diagnosed either on the basis of the symptoms of inattention or on the basis of the symptoms of hyperactivity-impulsivity.

Although the co-occurrence of behavioural problems has been demonstrated in many cases, investigators have started to differentiate between subjects who are aggressive, but not hyperactive,32,33 aggressive, but not inattentive,34 and aggressive, but not prosocial.35,36 Further, depressive subjects who experience symptoms of anxiety have been differentiated from subjects who express symptoms of depression only.26,57

In research on children’s behavioural and emotional problems, the question of situational specificity has often been raised. Typically, low correlations between different informants have been perceived as evidence of low validity of the measure. In the absence of definite criteria of what constitutes a valid measure, it is very important to obtain assessments from different informants. Achenbach, McConaughy, and Howell38 studied cross-informant agreement meta-analytically. Of 119 studies which were reviewed, 23 dealt with agreement between peers and teachers; the mean correlation was found to be 0.44. Other studies have also shown that concurrent validity between peer and teacher assessments is usually moderate.41,39 The mean correlation between teacher and parental assessments is lower, 0.27 in 26 studies.38 When parental assessments have been compared with those of other informants, the greatest agreement has been between mothers and fathers, followed by mothers and teachers.40,41 The poorest agreement is between children’s self-assessments and those by adults,42 with parents 0.25 in 11 studies and teachers 0.20 in 17 studies.38 The 119 studies reviewed by Achenbach et al38 did not include any in which the agreement between peer nominations and parental assessments was compared.

One aim of the present study was to investigate the agreement between different informants, i.e. to assess concurrent validity of peer nominations using teacher and parental assessments as criteria. It was expected that concurrent validity of peer nominations will be higher when teacher, rather than parental, assessments are used as criteria, because peers and teachers observe children’s behaviour in common school settings. Additional aims of this study were to evaluate the internal structure, reliability, and discriminative validity of the multidimensional inventory with data from twin children.

Compared with Achenbach’s Child Behavior Checklist (CBCL),43 the multidimensional inventory developed in the present study covers both behavioural and emotional problems included in the CBCL, but also both active and passive adjusted behaviour – a broader coverage of children’s social behaviour. According to the model in Figure1, which was used as a heuristic framework for the construction of the peer nomination inventory, the inventory includes four major components (TypesA to D). We expected them to constitute the internal structure of the inventory.

Discriminative validity was studied, first, in relation to gender differences. It was expected that TypeA behaviour would be more common among boys than girls,34 whereas girls were expected to exhibit TypeD behaviour,44,45 and TypesB and C behaviour46,47 more than boys. Second, discriminative validity was evaluated in the relative resemblance of monozygotic and dizygotic twins: larger intra-pair differences were expected among the dizygotic twin pairs, because research by others48–51 has found consistent evidence of genetic effects on children’s behavioural characteristics assessed by parental ratings on the CBCL. We extend these reports, using assessments from peers and teachers, as well as parents, on twin samples that are age-standardised and representative of both same and opposite sex DZ twin pairs.

Method

Participants

The reported data are from a current longitudinal study of behavioural development and health habits, FinnTwin12, which will ascertain and study about
2800 twin pairs from five consecutive and complete birth cohorts of 12-year-old Finnish twin children. Data here reported were collected on intensively-studied subsets of the first three of five nationwide birth cohorts. In 1994 to 1996, all Finnish families with twins born in 1983 to 1985 were identified from the nation's Central Population Registry as part of Finnish Twin Cohort studies. Fifty-eight percent of the identified families returned a questionnaire on the twins' childhood development and medical history. Permission to contact schools attended by the twin children, to obtain teacher and peer assessments, was obtained from 92.7% of these respondent families.

Peer nominations were collected for a sub-sample of twins (about 40% of all twin pairs in each cohort) selected for intensive follow-up. The selection was made randomly from families throughout the entire country for about half of the intensively studied subjects; the random subsample was enriched with an 'at risk' subsample of twins selected on the basis of their parents' scores on the Malmö-modified Michigan Alcoholism Screening Test, first used in Finland by Seppä.

The three-year sub-sample consisted of 477 twin girls and 467 twin boys, and these 944 twins were included in all analyses requiring neither zygosity confirmation nor data from the co-twin. Comparisons of MZ and DZ twins were limited to pairs of confirmed zygosity with complete data from both co-twins in each pair. Sample sizes for these comparisons were 154 MZ pairs, 132 same-sex DZ pairs, and 127 opposite-sex dizygotic twin-pairs. Zygosity was determined from twins' perceived similarity and confusability of appearance, as reported by the twins and their parents in mailed questionnaires; the questionnaire items have been validated against blood-typing. Zygosity could not be assigned from a questionnaire on n = 45 same-sex pairs (their zygosity awaits DNA confirmation), and MPNI data were missing or incomplete in 28 pairs (including nine pairs of known zygosity, 4 MZ and 5 DZ, in which one twin was in a special education class, prohibiting collection of peer nomination data).

Permission to conduct peer nominations at school was sought from school principals; permission was obtained from 99%. Teacher and parental assessments of twins' social behaviour, collected by mailed questionnaires, were available for all twins in the sub-sample. The twins were in 503 school classes; classroom teachers for each of the 503 classes participated in teacher assessments, and the classes included 12937 pupils who participated in peer nominations. Two-thirds of parental assessments were reported to have been filled in by the mother, and one-third by both parents jointly.

Measures and variables

The multidimensional inventory of children's social behaviour used in this study was developed for peer nomination (MPNI; the Multidimensional Peer Nomination Inventory). Parental and Teacher Rating Forms were developed from the MPNI. The MPNI included 30 items; seven items were added to Parental and Teacher Rating Forms for enriching coverage of Attention-Deficit/Hyperactivity Disorder (Inattention and Hyperactivity-Impulsivity), anxiety, and depression (see Appendix).

The core items of the MPNI were developed by Pulkkinen to represent a model of emotional and behavioural regulation. Items for assessment of four behavioural types - aggression, anxiety, constructiveness and compliance - were included in the inventory. Additional items originated from different studies: they concerned indirect aggression, internalising and externalising problems, and Attention-Deficit/Hyperactivity Disorder (DSM-IV, American Psychiatric Association, 1994). A series of pilot studies was carried out for item selection and for finding wording appropriate for the items, to adjust to the age of children, as suggested by Banbury and Wellington.

Peer nominations were made in classroom settings, under the supervision of a member of our research staff. Except for the first two items, MPNI items were presented in mixed order. Each pupil was given a pad with 32 pages. The first two items, 'Who are you?' and 'Which of your classmates are absent from school today?', provided practice with the demands of the method. All pages were similar, and included the forenames of all classmates, divided by gender and presented in alphabetical order. All MPNI questions (eg ‘Which of your classmates may hurt other kids when they’re angry, for instance, by hitting, kicking, or throwing things at them?') were read by the research assistant, and the pupils were asked to respond by crossing out names of those classmates who often engage in the behaviour described. Pupils were asked to nominate three female and three male classmates, if possible, who best fit the described behaviour of each MPNI item. Each pupil in the class received some number of nomination ‘votes’ for each item; these were expressed as a percentage, a ratio of received votes to the maximum number of votes the pupil could have received.

In the Parent Rating and Teacher Rating Forms of MPNI, the items were presented in the form of statements (eg ‘Hurts other kids when they’re angry...’). The teachers and parents were asked to rate each twin on every item on a four-point scale where 0 = does not apply, 1 = applies sometimes, but not consistently, 2 = certainly applies, but not in...
a pronounced way, and 3 = applies in a pronounced way.

Data analysis

The structure of the Multidimensional Inventory was studied using a factor analysis (PAF) and varimax rotation. An eigenvalue larger than one was used as a criterion to derive the scales. Varimax rotation was chosen because the components of social behaviour were orthogonally related in the model (Figure 1). Sub-components of the correlating variables of the major factors were studied with further factor analyses and oblimin (Kaiser normalisation) to extract possible specific components from a more homogeneous set of variables. Coefficients alpha were calculated to assess reliability of the scales formed on the basis of the factor analyses. The concurrent validity of the assessments was evaluated with correlation coefficients between peer nominations, teacher, and parental assessments on both scale and item level. The analyses were replicated with a random half sub-sample that included only one co-twin from each twin pair. Those results confirmed the findings reported here on the full sample. The discriminative validity of the inventory was studied by comparing, for each scale, the means of girls and boys, and the intra-pair correlations of MPNI ratings of MZ, same-sex DZ, and opposite-sex DZ twin pairs.

Results

Internal structure of the inventory

A factor analysis (principal axis method and varimax rotation) of the MPNI and the Teacher and Parent Forms of it resulted in three factors. The first factor was loaded by items for Hyperactivity-Impulsivity (‘Is restless’, ‘Does not wait for his turn’), Aggression (‘Hurts other children when angry’), and Inattention (‘Has poor concentration’). Compliance (eg ‘Is peaceable’) loaded negatively on this factor. The first factor describes poor control over behaviour which causes social problems by violating social order or other people’s rights, and was identified as Behavioral Problems. It explained 34.1% of the variance of peer nominations (PN), and 31.7% and 18.5% of the variance of teacher ratings (TR) and parental ratings (PR), respectively. Factor analyses with oblimin rotation on the items that loaded on the first factor did not reveal significant components for PN, but for TR and PR intercorrelating factors for Hyperactivity-Impulsivity, Aggression, and Inattention emerged; in PR, aggression also divided into direct Aggression and Indirect Aggression.

The second factor was loaded by Constructiveness, which reflects consideration of other people’s rights and needs (‘Tries to act reasonably’), Compliance, which reflects emotional balance (‘Is peaceable’), and Social Activity, including leadership, popularity, and interaction with other children. The second factor describes strong control of behaviour and emotions, and was identified as Adjustment. It explained 17.6% of the variance of PN, and 10.6% and 6.2% of the variance of TR and PR, respectively. Further factor analyses with oblimin rotation with the items that loaded on the second factor resulted in separate components for Constructiveness and Compliance for all informants (peers, teachers, and parents). A third component was Social Activity. A fourth component was formed by Helping Behaviour (‘Lends a helping hand’ and ‘Stands up for smaller and weaker peers’) for TR only.

The third factor was loaded by items for anxiety and depression. In PN, the highest loading was on the item ‘Is frightened and nervous about new things or new situations’, and in TR and PR on the item ‘Worries a lot’. Negative loadings were obtained on the items for Social Activity. The factor describes poor emotional control that renders one’s life difficult and causes personal problems. It was identified as Emotional Problems. Further factor analyses with oblimin rotation with the items that loaded on this factor showed that PN and TR were not divided into components, but two correlating components, Depression and Social Anxiety, emerged in PR.

An item for victimisation (‘Gets teased and taunted a lot’) was included in teacher and parental assessments. It was loaded on the factor for Emotional Problems. If it was included in a further factor analysis of PR with an oblimin rotation, it was not loaded on the Depression or Social Anxiety factors, but rather formed with loneliness a separate factor for Social Isolation.

In the Appendix, the items are grouped according to the components extracted; the items with the highest loadings on each factor are listed first (loadings higher than 0.30 were considered). The three major factors: Behavioural Problems, Emotional Problems, and Adjustment were very similar for PN, TR, and PR. They can be divided into sub-components as shown in the Appendix, if more specific scale measurement is needed.

Internal consistency

Scale reliability was studied using coefficients alpha (Table 1). The reliability of peer nominations was high, especially for Behavioural Problems and its sub-scales. The reliability of the three major scales was also satisfactory for teacher and parental assessments. Parental assessments of subscales were less
reliable for Social Anxiety and Compliance. Reliability coefficients of teacher assessments fell between those of peer nominations and parental assessments for most scales. In Emotional Problems, the overall level of reliability was lower than in the two other main scales, and differences between the informants in the size of reliability coefficients were smaller. Compared with boys, girls' depression and anxiety were less reliably rated by peers, although gender differences were generally small in the internal consistency of the scales.

Concurrent validity of ratings

Correlation coefficients between the corresponding items for peer nominations, teacher and parental assessments revealed that, for most items (25 out of 30), correlations were significantly higher between PN and TR (0.39 on average) than between PN and PR (0.19 on average) or between TR and PR (0.22 on average). The latter did not differ significantly from each other. The highest correlation (0.59, P < 0.001) between PN and TR was found for the item 'Is restless', and the lowest correlation (0.10, P < 0.01) for the item 'Withdraws from a difficult situation and starts doing something else', which correlated with Compliance in PN, but with Emotional Problems in TR.

The correlations between different informants for each scale are shown in Table 2. Correlations between PN and TR were higher than those between PN and PR or between TR and PR. Correlations were quite similar for boys and girls, although agreement was slightly higher in assessing boys' than girls' behaviour. Concurrent validity was highest for Behavioral Problems.

Discriminative validity

Gender differences To study the discriminative validity of the measures, gender differences in mean ratings were examined using t-test of independent samples (Table 3). Comparisons between girls and boys showed that boys scored higher than girls on the scales for Behavioral Problems, whereas girls scored higher than boys on Adjustment. For Emotional Problems, a slight gender difference existed only in peer nomination, Social Anxiety being higher in girls than in boys. For Depression, boys and girls did not differ.

Table 1: Reliability of the scales for peer nominations, teacher and parental assessments: coefficients alpha (477 girls, 467 boys)

<table>
<thead>
<tr>
<th>Scales</th>
<th>Peer nomination</th>
<th>Teacher assessment</th>
<th>Parental assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Girls</td>
<td>Boys</td>
<td>Girls</td>
</tr>
<tr>
<td>Behavioral Problems</td>
<td>0.95</td>
<td>0.97</td>
<td>0.93</td>
</tr>
<tr>
<td>Hyperactivity-Impulsivity</td>
<td>0.91</td>
<td>0.96</td>
<td>0.90</td>
</tr>
<tr>
<td>Aggression</td>
<td>0.92</td>
<td>0.94</td>
<td>0.85</td>
</tr>
<tr>
<td>Inattention</td>
<td>0.92</td>
<td>0.95</td>
<td>0.89</td>
</tr>
<tr>
<td>Adjustment</td>
<td>0.91</td>
<td>0.91</td>
<td>0.85</td>
</tr>
<tr>
<td>Constructiveness</td>
<td>0.93</td>
<td>0.93</td>
<td>0.82</td>
</tr>
<tr>
<td>Compliance</td>
<td>0.79</td>
<td>0.79</td>
<td>0.67</td>
</tr>
<tr>
<td>Social Activity</td>
<td>0.82</td>
<td>0.85</td>
<td>0.67</td>
</tr>
<tr>
<td>Emotional Problems</td>
<td>0.75</td>
<td>0.87</td>
<td>0.77</td>
</tr>
<tr>
<td>Depression</td>
<td>0.66</td>
<td>0.79</td>
<td>0.70</td>
</tr>
<tr>
<td>Social Anxiety</td>
<td>0.68</td>
<td>0.81</td>
<td>0.69</td>
</tr>
</tbody>
</table>

See Appendix. All correlations are significant at P<0.001 level except Parent/Peer correlation for Compliance in girls (P<0.01).

Table 2: Correlations between peer nominations, teacher parental assessments (477 girls and 467 boys) for each scale

<table>
<thead>
<tr>
<th>Scales</th>
<th>Peer/Teacher</th>
<th>Teacher/Parent</th>
<th>Parent/Peer</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All</td>
<td>Girls</td>
<td>Boys</td>
</tr>
<tr>
<td>Behavioral Problems</td>
<td>0.62a</td>
<td>0.58</td>
<td>0.65c</td>
</tr>
<tr>
<td>Hyperactivity-Impulsivity</td>
<td>0.58a</td>
<td>0.55</td>
<td>0.63c</td>
</tr>
<tr>
<td>Aggression</td>
<td>0.56a</td>
<td>0.54</td>
<td>0.56</td>
</tr>
<tr>
<td>Inattention</td>
<td>0.52a</td>
<td>0.45</td>
<td>0.57c</td>
</tr>
<tr>
<td>Adjustment</td>
<td>0.52a</td>
<td>0.48</td>
<td>0.54</td>
</tr>
<tr>
<td>Constructiveness</td>
<td>0.51a</td>
<td>0.46</td>
<td>0.50</td>
</tr>
<tr>
<td>Compliance</td>
<td>0.48a</td>
<td>0.36</td>
<td>0.49</td>
</tr>
<tr>
<td>Social Activity</td>
<td>0.55a</td>
<td>0.48</td>
<td>0.58</td>
</tr>
<tr>
<td>Emotional Problems</td>
<td>0.34a</td>
<td>0.36</td>
<td>0.33</td>
</tr>
<tr>
<td>Depression</td>
<td>0.30a</td>
<td>0.29</td>
<td>0.33</td>
</tr>
<tr>
<td>Social Anxiety</td>
<td>0.32a</td>
<td>0.37</td>
<td>0.32</td>
</tr>
</tbody>
</table>
Comparison between monozygotic and dizygotic twins. In a further test of discriminative validity of the MPNI, intra-pair correlations between monozygotic (MZ), same-sex dizygotic (SSDZ), and opposite-sex dizygotic (OSDZ) twins were calculated (Table 4). Consistently, across scales and across informants, correlations were higher for MZ co-twins, with less variation across assessors in MZ correlations than in those for DZ co-twins. But across informants, significant similarity was also found for both SSDZ and OSDZ twins; peer nominations and teacher ratings yielded significant correlations for DZ twins across all MPNI scales, and with but two exceptions, that was also true for correlations derived from parental ratings. Absolute intra-pair difference scores were compared for MZ and DZ twin pairs, as other researchers (e.g., Goodman and Stevenson48) have reported greater intra-pair differences in DZ co-twins in parental and teacher assessments of social behaviours; we replicated those results with peer nominations on the MPNI, as additional evidence of its discriminative validity. Compared with MZ twin pairs, absolute intra-pair differences were greater for both same-sex and opposite-sex DZ twins, across all scales and across all three sets of raters.

### Table 3 Gender differences in the means of the scales for peer nominations, teacher assessments and parental assessments; t-test

<table>
<thead>
<tr>
<th>Scales</th>
<th>Peer nomination</th>
<th>Teacher assessment</th>
<th>Parental assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Girls Boys</td>
<td>Girls Boys</td>
<td>Girls Boys</td>
</tr>
<tr>
<td><strong>n</strong></td>
<td>477 467</td>
<td>477 467</td>
<td>477 467</td>
</tr>
<tr>
<td>Behavioral Problems</td>
<td>M 12.3 18.6***</td>
<td>0.49 0.88***</td>
<td>0.60 0.79***</td>
</tr>
<tr>
<td>SD 13.8 20.9</td>
<td>0.48 0.65</td>
<td>0.35 0.41</td>
<td></td>
</tr>
<tr>
<td>Hyperactivity-Impulsivity</td>
<td>M 14.2 20.5***</td>
<td>0.45 0.92***</td>
<td>0.62 0.85***</td>
</tr>
<tr>
<td>SD 16.8 24.2</td>
<td>0.56 0.81</td>
<td>0.44 0.56</td>
<td></td>
</tr>
<tr>
<td>Aggression</td>
<td>M 12.1 17.0***</td>
<td>0.53 0.79***</td>
<td>0.57 0.65***</td>
</tr>
<tr>
<td>SD 12.6 18.0</td>
<td>0.57 0.68</td>
<td>0.38 0.42</td>
<td></td>
</tr>
<tr>
<td>Inattention</td>
<td>M 10.5 18.2***</td>
<td>0.48 0.93***</td>
<td>0.60 0.85***</td>
</tr>
<tr>
<td>SD 15.6 24.2</td>
<td>0.48 0.45</td>
<td>0.46 0.51</td>
<td></td>
</tr>
<tr>
<td>Adjustment</td>
<td>M 20.0* 18.1</td>
<td>1.97*** 1.69</td>
<td>2.11*** 1.95</td>
</tr>
<tr>
<td>SD 12.4 12.4</td>
<td>0.50 0.57</td>
<td>0.37 0.39</td>
<td></td>
</tr>
<tr>
<td>Constructiveness</td>
<td>M 19.7*** 15.8</td>
<td>1.94*** 1.62</td>
<td>2.22*** 2.01</td>
</tr>
<tr>
<td>SD 15.5 15.5</td>
<td>0.63 0.66</td>
<td>0.49 0.52</td>
<td></td>
</tr>
<tr>
<td>Compliance</td>
<td>M 17.8* 15.9</td>
<td>2.07*** 1.72</td>
<td>1.90*** 1.73</td>
</tr>
<tr>
<td>SD 13.8 13.9</td>
<td>0.63 0.73</td>
<td>0.48 0.52</td>
<td></td>
</tr>
<tr>
<td>Social Activity</td>
<td>M 22.5 22.5</td>
<td>1.92*** 1.75</td>
<td>2.23*** 2.13</td>
</tr>
<tr>
<td>SD 16.2 17.8</td>
<td>0.64 0.67</td>
<td>0.51 0.54</td>
<td></td>
</tr>
<tr>
<td>Emotional Problems</td>
<td>M 10.9* 9.5</td>
<td>0.79 0.75</td>
<td>0.78 0.78</td>
</tr>
<tr>
<td>SD 9.3 11.0</td>
<td>0.54 0.59</td>
<td>0.42 0.43</td>
<td></td>
</tr>
<tr>
<td>Depression</td>
<td>M 10.4 9.5</td>
<td>0.67 0.67</td>
<td>0.75 0.76</td>
</tr>
<tr>
<td>SD 8.8 10.9</td>
<td>0.50 0.55</td>
<td>0.43 0.42</td>
<td></td>
</tr>
<tr>
<td>Social Anxiety</td>
<td>M 11.4* 9.5</td>
<td>0.91 0.83</td>
<td>0.82 0.80</td>
</tr>
<tr>
<td>SD 12.4 12.5</td>
<td>0.74 0.75</td>
<td>0.58 0.58</td>
<td></td>
</tr>
</tbody>
</table>

* P<0.05, ** P<0.01, *** P<0.001.

### Table 4 Correlations between monozygotic (MZ), same-sex dizygotic (SSDZ), and opposite-sex dizygotic (OSDZ) co-twins

<table>
<thead>
<tr>
<th>Scales</th>
<th>Peer nomination1</th>
<th>Teacher assessment2</th>
<th>Parental assessment2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MZ OSDZ SSDZ</td>
<td>MZ OSDZ SSDZ MZ</td>
<td></td>
</tr>
<tr>
<td>Behavioral Problems</td>
<td>0.32 0.48 0.72***</td>
<td>0.50 0.60 0.84***</td>
<td>0.33 0.39 0.77***</td>
</tr>
<tr>
<td>Hyperactivity-Impulsivity</td>
<td>0.26 0.44 0.70***</td>
<td>0.54 0.59 0.87***</td>
<td>0.21 0.11 0.77***</td>
</tr>
<tr>
<td>Aggression</td>
<td>0.33 0.51 0.76***</td>
<td>0.46 0.59 0.82***</td>
<td>0.53 0.55 0.74***</td>
</tr>
<tr>
<td>Inattention</td>
<td>0.25 0.44 0.71***</td>
<td>0.53 0.61 0.79***</td>
<td>0.22 0.38 0.68***</td>
</tr>
<tr>
<td>Adjustment</td>
<td>0.42 0.37 0.79***</td>
<td>0.44 0.60 0.82***</td>
<td>0.37 0.46 0.83***</td>
</tr>
<tr>
<td>Constructiveness</td>
<td>0.42 0.39 0.78***</td>
<td>0.53 0.59 0.85***</td>
<td>0.42 0.38 0.81***</td>
</tr>
<tr>
<td>Compliance</td>
<td>0.38 0.42 0.70***</td>
<td>0.38 0.57 0.82***</td>
<td>0.22 0.20 0.68***</td>
</tr>
<tr>
<td>Social Activity</td>
<td>0.53 0.44 0.82***</td>
<td>0.38 0.50 0.76***</td>
<td>0.25 0.26 0.61***</td>
</tr>
<tr>
<td>Emotional Problems</td>
<td>0.51 0.41 0.79***</td>
<td>0.44 0.52 0.69***</td>
<td>0.33 0.29 0.59***</td>
</tr>
<tr>
<td>Depression</td>
<td>0.42 0.39 0.77***</td>
<td>0.49 0.54 0.57</td>
<td>0.36 0.42 0.43</td>
</tr>
<tr>
<td>Social Anxiety</td>
<td>0.48 0.39 0.78***</td>
<td>0.41 0.51 0.80***</td>
<td>0.24 0.12 0.64***</td>
</tr>
</tbody>
</table>

Difference between MZ and DZ correlations: *** P<0.001. ** P<0.01. *=non significant correlation.

1Spearman rank order correlations.
2Product moment correlations.
Discussion

Results confirm that the multidimensional inventory developed for peer nomination, and teacher and parental assessments is reliable and has both concurrent and discriminative validity. The internal structure of the MPNI, MPNI Teacher Form, and MPNI Parental Form, could be described in terms of three major dimensions: Behavioral Problems, Emotional Problems, and Adjustment. Except for the Q-sort, available behavioural checklists or inventories concern either behavioural or emotional problems, or both, such as the widely used instrument by Achenbach. Many instruments, such as the Behavior Assessment System for Children are complex by content, and not usable for peer nomination.

The reliability was very high for Behavioral Problems, which means that there was a redundancy in items; a smaller number of items would be sufficient for the reliable assessment of behavioural problems. The reliability of the scale for Emotional Problems showed that children can indeed identify symptoms of anxiety and depression in their classmates, as Ezpeleta et al have argued.

The discriminative validity of the multidimensional inventory was good. As expected, peer nominations differentiated between boys and girls: boys were rated behaviourally more undercontrolled than girls, as found by McGee et al, and girls' behaviour was seen as more controlled, replicating Graziano et al. A difference in Emotional Problems existed only in Social Anxiety with girls scoring higher than boys; no differences were found for assessed Depression. Several studies have shown that adolescent girls are more depressed than boys, but in pre-adolescence, findings are inconsistent. Nolen-Hoeksema and Girgus assume that around ages 13–14, gender differences in depression emerge. Participants in the present study were 11–12 years old.

It was expected that the internal structure of the inventory could be described in terms of four components, but three major factors were extracted. For testing a circumplex model as described in Figure 1, a factor analysis is not the best method. A correspondence analysis would fit better with the data. Also, two unrotated principal axis produce a two-dimensional structure, where all four components (Types A to D) are defined by the factors for emotional and behavioural regulation. However, two factors explain only part of the variance. When the number of factors is increased using an iterative principal axis method and varimax rotation, factors are formed by the groups of items which share common variance. Within the set of variables of the inventory where the number of items for problem behaviour were enriched, major groups of common variance were related to behavioural problems (Type A), emotional problems (Type D), and adjustment (Types B and C combined). Results show that discrimination in children's behaviour between active and passive problem behaviour, in contrast to adjusted behaviour, is most relevant. The two types of adjusted behaviour emerged, however, as separate components when a factor analysis was made with a more homogeneous set of variables for adjusted behaviour.

In person-oriented analyses using a clustering technique, three clusters for personality types have similarly been identified from Q-sorts, the Big Five factors of personality, and personality style variables with a sample of children and adults. The three clusters have been labelled differently by different investigators, eg undercontrolled or conflicted; overcontrolled or introverted; and adjusted or resilient individuals. Overcontrolled children in the study by Robins et al displayed both anxious and compliant behaviour, whereas Resilient children exhibited socially active, adjusted behaviour. The clusters (and factors) extracted depend on the variables included in the analysis. If all four components of the model in Figure 1 are balanced among the variables, clusters for each of them will also be extracted. Longitudinal study confirms the relevance for developmental processes of discriminating between children’s active and passive adjusted behaviours.

The analysis of the sub-components of the major dimensions with further factor analyses and oblimin rotations revealed that the smallest number of sub-components emerged in peer nomination. The highest number of meaningful sub-components emerged in parental ratings. This fact suggests that the parents make accurate discriminations while observing their children's behaviour. The reliability of parental assessment was, however, lower than peer nominations and teacher assessment, and parental assessments correlated less highly with peer nominations and teacher assessments than did peer nominations and teacher assessments with each other. The mean correlation between peer nominations and teacher assessments was about the same (0.48) as Achenbach et al found in their meta-analysis for cross-informant agreement between peers and teachers (0.44). The corresponding figure in the present study for the agreement between teacher and parental assessments was 0.35 (Achenbach et al reported 0.27) and between peer nominations and parental findings 0.27. The correlations obtained were slightly higher than the mean correlations in the Achenbach et al meta-analysis. It was not affected by the twin sample; the correlations for half a sample including only one of the co-twins from each pair were in the corresponding order: 0.50, 0.34, and 0.28. The agreement was higher (P < 0.001) between
peers and teachers than between the other informants, and it was higher (P < 0.05) between parents and teachers than between parents and peers. Classmates, teachers, and parents observe children in different situations and from different, adult or child, perspectives. Children may also behave differently in school and home.68

The pattern of MZ and DZ correlations was remarkably consistent, across MPNI scales, particularly for peer nominations, where all MZ correlations were ≥ 0.70, and DZ correlations ranged from 0.29 to 0.53. Across informants, MZ correlations were consistently high, approximating a doubling of the DZ correlations, suggesting that additive genetic influences play a major role in determining inter-individual variation in behavioural and emotional problems of pre-teenage children. Significant behavioural similarity of DZ co-twins was found in peer nominations and in ratings by both teachers and parents, with little or no evidence of contrast effects that produce unexpectedly low DZ correlations, as frequently found by others (eg Saudino and Cherry69). Finally, the significance and similarity of same-sex and opposite-sex DZ twin correlations suggests, that at this age, no major gender differences in the sources and components of behavioural variance assessed by the MPNI exist. When FinnTwin12 data collection is complete, two additional twin cohorts will be enrolled into the study, and formal genetic analysis of sex-limitation effects will be undertaken.

Acknowledgements

Data collection was supported by the National Institute on Alcohol Abuse and Alcoholism (AA-09203 and AA-00145) and by the Academy of Finland. During a pilot study conducted during 1992–1993, Hanna Järventaus (f.Isomäki) and Päivi Kukkonen participated in the development of the Multidimensional Peer Nomination Inventory in their Master’s theses. Elina Läätä worked for the project in 1994–1995, participating in data collection and tentative analyses of the MPNI data. Ari Mäkiaho and Olli Louhela have assisted in data management and analysis.

References

Appendix:

Behavioral problems

Hyperactivity-impulsivity

1 Which of your classmates are restless, unable to sit in class? (4)
2 Teacher and parental ratings only: Is hyperactive. (37)
3 Which of your classmates are the most talkative? (15)
4 Which of your classmates seem to have difficulties in waiting for their own turn? (29)
5 Teacher and parental ratings: Is disobedient at school/home. (33)
6 Which of your classmates often act rashly, i.e. without thinking about the possible consequences? (11)
7 Teacher and parental ratings: Runs about and climbs everywhere in spite of warnings. (31)

Aggression

Direct Aggression:

8 Which of your classmates may hurt other kids when they are angry, e.g. by hitting, kicking, or throwing things at them? (21)
9 Which of your classmates call people names when they are angry with them? (27)
10 Which of your classmates tease other kids or attack them for no reason at all? (13)
11 Which of your classmates bully smaller and weaker kids? (25)

Indirect Aggression:

12 Which of your classmates go round telling people's secrets to others? (18)
13 Which of your classmates try to leave other kids out of the company of others eg by saying: 'Let's not play/hang around with him/her'? (9)

Inattention

14 Teacher and parental ratings: Is conscientious with homework. (Reversed) (34)
15 Teacher and parental ratings: Is forgetful. (36)
16 Which of your classmates tend to ignore instructions? (17)
17 Which of your classmates cannot concentrate on anything? (8)

Adjustment

Constructiveness

18 Which of your classmates try to act reasonably even in difficult situations? (7)
19 Which of your classmates are considered reliable by all kids in your class? (20)
20 Which of your classmates are able to sort things out by talking? (12)

Helping Behavior:

21 Which of your classmates are ready to lend a helping hand when they see someone in need of that? (26)
22 Which of your classmates often stand up for smaller and weaker kids? (16)

Compliance

23 Which of your classmates are peaceable and have a lot of patience? (5)
24 Which of your classmates never quarrel with others? (28)
25 Which of your classmates are often kind and friendly to others (2)
26 Peer nominations only: Which of your classmates withdraw from difficult situations and start doing something else instead? (23)

Social Activity

27 Which of your classmates would you invite to your birthday party? (In teacher and parental rating: Is popular among his/her classmates.) (30)
28 Imagine your class going on an outing. Which of your classmates do you think would make a good leader? (Teacher rating: Is a good leader and would be suitable to lead a class outing; Parental rating: When you have a group of children visiting, your child would be good at leading group play.) (1)

29 Which of your classmates are often with other kids during recess and after school? (22)

Emotional problems
Depression

30 Which of your classmates worry a lot? (24)

31 Which of your classmates seem to be sad and depressed a lot of the time? (3)

32 Which of your classmates are easily offended/ start crying if someone is nasty to them? (6)

33 Which of your classmates are lonesome, without friends? (14)

34 Teacher and parental ratings only: Clings to adults or is too dependent. (32)

Social Anxiety

35 Which of your classmates are shy with other kids? (10)

36 Which of your classmates are frightened and nervous about new things or new situations? (19)

Victimisation

37 Teacher and parental ratings only: Gets teased and taunted a lot. (35)