Chapter 18

CROSS-INFORMANT AGREEMENT AND TEACHER NOMINATION TECHNIQUE IN THE ASSESSMENT OF CHILDREN BEHAVIOR PROBLEMS

Helena Klimusová¹, Iva Burešová¹, & Ivo Čermák²

¹Institute of Psychology, Faculty of Arts, Masaryk University, Czech Republic
²Institute of Psychology, The Academy of Sciences of the Czech Republic, Czech Republic

ABSTRACT

The aim of our study was to (1) examine the degree of agreement among children, parents and teachers on the scales and items of the Czech version of the Child Behavior Checklist, and (2) to assess a teacher nomination technique as a short and simple tool for the recognition of children with higher levels of self-reported problems. The first part of the study was conducted on a sample of 300 children (aged 11-16). The overall degree of cross-informant agreement was low, particularly for ratings of boys. The highest correlations were found between the ratings of parents and teachers (median correlation 0.336), followed by the correlations between the ratings of parents and children (median correlation 0.316). The agreement between teachers and children’s ratings was generally poor (median correlation 0.115), with the exception of moderate agreement on the Externalizing Behavior scales. The nomination technique was mainly based on the TRF scales and contained nine short descriptions; the sample consisted of 145 children from non-clinical population (aged 11-12). Only a low proportion of children were named by the teachers at each problem child description. Four significant associations were found between the teacher's ratings in the nomination technique and the categorical ratings based on children self-reports (YSR). Results of both parts of the study are consistent with previous findings and emphasize the importance of utilizing multiple sources in the assessment of children behavior problems.

Keywords: CBCL, informant’s agreement, nomination technique, teachers, child problem behavior.

1. INTRODUCTION

In the past, the assessment of a child's behavior problems frequently relied on a single informant, typically a parent. However, assessment methods have increasingly demanded the utilization of multiple sources of information, including the children themselves. Multiple informants may each contribute unique information about the child's behavior: Parents may not be aware of their child's intrapsychic symptoms or concealed conduct problems and they may only be able to observe their child in a limited range of situations. Children's behavior varies from one context and interaction partner to another; and many child problems are not consistently present across different settings but may occur exclusively at home or at school (Cantwell, Lewinsohn, Rohde, & Seeley, 1997; Achenbach & McConaughy, 1997).

2. BACKGROUND

Achenbach, McConaughy, & Howell (1987) conducted a meta-analysis of studies that used various assessment instruments and found only moderate correlations between different informants’ ratings of child behavior problems. Poor concordance between parent ratings, peer ratings and self-ratings of social behavior – the Social Competence section of the CBCL and related instruments – has also been reported (Schneider & Byrne, 1989). However, Achenbach et al. (1987) do not equate the low cross-informant agreement in child assessment with unreliability; instead, they see it as reflecting the cross-situational specificity of the children’s emotional and behavioral problems.

Therefore, different sources of information do not always provide consistent reports. Proper choice of informants for particular categories of child problems may lead to improved
diagnostic and predictive validity of the assessment. The reports of different informants may differ in terms of reliability (i.e., test-retest stability) and predictive power; certain informants may be superior to others in assessing specific child behaviors. Loeber, Green, Lahey, and Stouthamer-Loeber (1990) surveyed mental health professionals’ perceptions of relative usefulness of children, their mothers and teachers as informants on children’s problem behavior. Children and their mothers were perceived as more useful informants than teachers on children’s internalizing behavior, teachers were seen as the most useful informants on children hyperactivity and attention problems, and children were perceived as the least useful informants on their own attention and hyperactivity problems and oppositional behavior. The study of Loeber and colleagues (1990) has been limited to comparisons involving 7-to-12-year-old children. The authors suppose that the usefulness of informants changes from late childhood to adolescence, expecting adolescents to be the best informants on their own internalizing problems and on concealed conduct problems. Phares (1997) examined mothers’ and fathers’ opinions on the accuracy of various types of informants (mothers, fathers, teachers, peers, and children themselves) of children’s emotional and behavioral problems. Mothers were both by themselves and by fathers perceived as the best informants on children’s internalizing problems; mothers and teachers were seen as the most accurate in reporting externalizing behavior; and mothers, fathers, and children themselves were perceived as superior in reporting family problems.

Several studies identified teacher ratings of children problems as good predictors, for example, of referral for mental health services. The ability of parents, teachers, and self-reports to predict symptoms of maladjustment in 11-to-14-year-old children over a 4-year interval were examined in the study by Verhulst, Dekker, and Van der Ende (1997). Each of the three informants made a unique contribution to the prediction of maladjustment. Factors that predicted actual referral and parental need for professional help were different from the factors that predicted children’s own perceptions of having problems or of needing professional help. Surprisingly, teachers’ ratings of internalizing problems, often viewed as less accurate compared to those of parents, were found to be highly useful in predicting the child’s own perceptions of having problems. This finding is consistent with the results of another longitudinal study (Mesman & Koot, 2000), which attempted to identify parent- and teacher-reported behaviors indicating the presence of child-reported depression and anxiety. Results showed that child-perceived depression and anxiety were only slightly related to parental ratings of the child’s problems, while they were moderately related to teacher-reported problems. The authors inferred that teachers were more likely than parents to notice internalizing problems and related problems in children reporting depression or anxiety. These findings are contradictory to those of Crowley and colleagues (1992), who examined the relationship between self-report, peer-report, and teacher-report measures of childhood depression on item level; they concluded that those three types of measures yielded scores that were primarily independent and thus measured generally uncorrelated constructs. These contradictory findings suggest that it is essential to include both children and adults as informants on internalizing problems.

Research studies have shown that adult-child agreement on externalizing behaviors is lower than adult-adult agreement on the same behavior (e.g., Edelbrock, Costello, Duclan, Calabro Conover, & Kalas, 1986). Moreover, the study of Stanger and Lewis (1993) on agreement among parents, teachers, and children identified teacher ratings of externalizing problems as the best predictors of referral for mental health services. Children may underreport conduct problems because they forget about them, especially when the reference period is long, or because they do not identify the behavior as problematical. As children grow older, their ability to assess their own behavior improves and their assessment becomes more concordant with that of adults (Edelbrock et al., 1985, 1986). On the other hand, children’s ability to hide problem behaviors (e.g., theft, drinking, or drug use) improves with increasing age as well. Achenbach et al. (1987) in their meta-analysis found that concordance between adults and children’s ratings decreased slightly from late childhood to adolescence. There is also some evidence that the more antisocial children are, the lower are their ratings of their own problematic behavior, compared with the assessment of others (Loeber et al., 1990; Sawyer,
For aggressive and anti-social behavior, assessment by peers might be the most accurate, even at an early age. In their longitudinal study, Clemans, Musci, Leoutsakos, and Ialongo (2014) compared the predictive validity of teacher, parent, and peer reports of aggressive behavior in the first grade for maladaptive outcomes (such as antisocial personality, substance use, incarceration history, risky sexual behavior, and failure to graduate from high school on time) in late adolescence and early adulthood. Peer reports were found to be the most accurate predictors of all outcomes.

Research studies indicate that teacher reports may show high accuracy and predictive validity for some types of problem behavior. However, long inventories, questionnaires, or even short screening tools might be very demanding and time consuming when the teacher is asked to assess a high number of students (often all students in the class). There has been an effort to simplify the teacher’s assessment in those cases. Understandably, the accuracy and validity of such simplified assessments has been examined. One of the more promising methods is the nomination technique: One or more short descriptions of problem behavior are presented to the teachers, who are asked to nominate (name) children who in their opinion match the description.

One of the first studies of the teacher nomination method for identifying child behavior problems was conducted by Green, Beck, Forehand, and Vosk (1980). The nomination captured conduct vs. withdrawal problems (the rest of the children were classified as normal). The results were validated against sociometric ratings by peers, behavioral observations and academic achievement measures. The children identified by teachers as having conduct problems differed from normal children in all of these criteria, whereas the children identified as withdrawn differed from normal pupils only in sociometric ratings and academic achievement scores. More importantly, the groups of children with conduct problems and withdrawal problems did not differ from each other significantly in any of the measures. These findings lead the authors of the study to the conclusion that teachers are able to recognize a child with a problem but they might find it difficult to identify the type of the problem.

Henry, Miller-Johnson, Simon, Schoeny, and The Multisite Violence Prevention Project (2007) used the teacher nomination method to identify socially influential, aggressive adolescents for participation in a violence prevention program. They compared the teacher’s nominations with peer nominations and found significant correlations between them. Dwyer, Nicholson, and Battistutta (2006) used a very simple form of the nomination technique: a single question asking if either parents or the child’s teacher believed the child was at a higher-than-average risk of developing a mental health problem in the future. They tested the predictive validity of this method as well as the more extensive Family Risk Factor Checklist (FRFC) for children’s internalizing, externalizing, and total behavior problems scores after one year. Both parents’ and teachers’ predictions were more valid for externalizing than internalizing problems, with teacher nominations showing greater predictive validity than parent nominations. The teacher nomination method had higher sensitivity for predicting all three scores than the FRFC had.

3. OBJECTIVES

The first purpose of our study was to examine the degree of agreement between children, parents and teachers on the scales and items of the Czech version of the Child Behavior Checklist. The instrument was translated into the Czech language and introduced in the Czech Republic by Ivo Čermák and it has been used mainly in research studies so far (e.g. Čermák & Urbánek, 1998; Čermák & Klimusová, 2000). In the present study, we decided to examine the agreement between categorical as well as continuous ratings of emotional and behavioral problems. This dual approach is recommended, for example, by Stanger and Lewis (1993). The second objective was to assess a short instrument of our own construction – a teacher nomination technique based on TRF scale descriptions. Our aim was to explore whether teachers are able to identify children with higher levels of self-reported problems by means of a short and simple instrument.
4. METHODS

The Czech versions of the Child Behavior Checklist (CBCL), the Youth Self-Report (YSR) and the Teacher Report Form (TRF) were administered to a sample of 300 young adolescents (aged 11-16 years; $M = 13.41; SD = 0.95; 58 \%$ boys). The correlations between informants' ratings were analyzed, as well as the patterns of conditional agreement among informants (Kappa coefficients).

The nomination technique consisted of nine short descriptions, and the teachers were to name children who matched each description. They could match any number of children to each description, and every child could be associated with any number of descriptions. The descriptions were to a large extent based on the items of TRF scales. For example, the description based on the Anxious/Depressed scale was as follows: *Child B often fears that he/she might have done something wrong; he/she seems to be unhappy, sad or depressed. He/she gives the impression that he/she feels guilty or inferior or that he/she feels that nobody likes him/her.* The sample consisted of 145 children of non-clinical population (11 and 12 years of age; $M = 11.17; SD = 0.38; 41\%$ boys). Since parent ratings were not available for this sub-sample, only comparisons with self-reports were possible.

5. RESULTS

5.1. CBCL scales cross-informant agreement

The Pearson correlation coefficients reflecting the agreement between parents’ and children’s reports, children’s and teachers’ reports, and parents’ and teachers’ reports on behavior problem scales are presented in Figure 1. The highest correlations were found between the ratings of parents and teachers (median correlation 0.336), followed by the correlations between the ratings of parents and children (median correlation 0.316). Parents’ rank orders of children agreed with those of teachers mainly on Somatic Complaints, Attention Problems, and the scales belonging to the Externalizing Behavior scale – Delinquent Behavior and Aggressive Behavior scales. The agreement between parents’ and children’s ratings was lowest on the Social Problems scale; the correlations for the rest of the scales varied slightly around $r = 0.350$. The agreement between teachers’ and children’s ratings was generally poor; the median correlation was 0.115 and only four of the eleven correlations were significant. Correlation coefficients higher than 0.200 were found only for the Externalizing Behavior scales.

The cross-informant agreement was generally better for assessment of girls than of boys. Regarding the agreement between parents and children’s ratings, the median correlations for the assessment of girls and the assessment of boys were 0.366 and 0.246, respectively. The median correlations between parents and teachers’ ratings were 0.346 for girls and 0.307 for boys. The lowest agreement was found between the children and teachers’ ratings of boys (median correlation 0.078); the median correlation for ratings of girls was somewhat higher (0.257). A very similar pattern of cross-informant agreement can be seen in Table 1 presenting the Kappa coefficients. The Kappas indicated that the degree of agreement in classifying the child among the 15\% highest scoring children by particular informants was low. The highest agreement was obtained for parents-children ratings (median Kappa of 0.240) and for parents-teachers ratings (median Kappa 0.186).
Table 1. Kappa coefficients for the CBCL, TRF, and YSR scales.

<table>
<thead>
<tr>
<th></th>
<th>Parents-Children (N=201)</th>
<th>Teachers-Children (N=238)</th>
<th>Parents-Teachers (N=293)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CBCL/YSR/TRF narrowband scales</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1) Withdrawn</td>
<td>0.111</td>
<td>0.026</td>
<td>0.049</td>
</tr>
<tr>
<td>(2) Somatic complaints</td>
<td>0.259***</td>
<td>0.003</td>
<td>0.291**</td>
</tr>
<tr>
<td>(3) Anxious/depressed</td>
<td>0.196**</td>
<td>0.026</td>
<td>0.113</td>
</tr>
<tr>
<td>(4) Social problems</td>
<td>0.010</td>
<td>0.074</td>
<td>0.216**</td>
</tr>
<tr>
<td>(5) Thought problems</td>
<td>0.162*</td>
<td>0.073</td>
<td>0.176*</td>
</tr>
<tr>
<td>(6) Attention problems</td>
<td>0.026</td>
<td>0.039</td>
<td>0.186*</td>
</tr>
<tr>
<td>(7) Delinquent behavior</td>
<td>0.296**</td>
<td>0.051</td>
<td>0.304**</td>
</tr>
<tr>
<td>(8) Aggressive behavior</td>
<td>0.305**</td>
<td>0.263**</td>
<td>0.123</td>
</tr>
<tr>
<td><strong>CBCL/YSR/TRF broadband scales</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Internalizing behavior</td>
<td>0.240**</td>
<td>0.064</td>
<td>0.193**</td>
</tr>
<tr>
<td>Externalizing behavior</td>
<td>0.266**</td>
<td>0.226**</td>
<td>0.132</td>
</tr>
<tr>
<td>Total problem score</td>
<td>0.240**</td>
<td>0.078</td>
<td>0.273**</td>
</tr>
</tbody>
</table>

*Note: *Kappa coefficient is significant at the 0.05 level
**Kappa coefficient is significant at the 0.01 level

Figure 1. Correlation coefficients for the CBCL/YSR, TRF/YSR, and CBCL/TRF scales.

5.2. Nomination technique

The teachers generally named only a low proportion of children at each child problem description. Most of the children (75.2 %) were not ascribed to any description. The ‘Aggressive Child’ description was most frequently ascribed, followed by the ‘Child with Social
Problems’ description. The Kappa coefficients were then computed as coefficients of agreement between the dichotomous variables based on the YSR scales (indicating whether a score did or did not belong to the 15% of cases with highest scores) and the presence or absence at a particular description. Four significant associations were found between the teacher’s ratings in the nomination technique and the categorical ratings based on the YSR. The Kappa of 0.171 ($p < 0.01$) was obtained for Description B and the Anxious/Depressed scale of the YSR: out of the five children marked as depressed by the teacher, three children scored above the 85th percentile in the self-reported Anxious/Depressed scale. The rest of the significant Kappas were found for the ratings in Description E – Aggressive Child, which was associated with the Aggressive Behavior scale ($\kappa = 0.232, p < 0.01$), Delinquent Behavior scale ($\kappa = 0.162, p < 0.01$), and the broadband Externalizing Behavior scale ($\kappa = 0.194, p < 0.01$). For all three associations, four of the twelve children nominated by the teacher scored above the 85th percentile in the self-report scales.

A current or past referral to counseling or child psychiatrist was reported by teachers for 15 children. It was associated with high self-rated scores in five narrowband scales: Withdrawn ($\kappa = 0.178, p < 0.01$), Thought Problems ($\kappa = 0.217, p < 0.01$), Attention Problems ($\kappa = 0.180, p < 0.01$), Aggressive Behavior ($\kappa = 0.190, p < 0.01$), and Delinquent Behavior ($\kappa = 0.251, p < 0.01$). The referral status was also associated with the broadband Externalizing Behavior scale ($\kappa = 0.228, p < 0.01$) and the Total Problem score ($\kappa = 0.223, p < 0.01$).

6. FUTURE RESEARCH DIRECTIONS

To further assess the utility of the nomination technique as a potential screening instrument, the measure should be administered concurrently with the TRF – the method on which the nomination descriptions were based. Hepburn et al. (2008) chose this approach in their pilot study of the nomination technique as a screening method for autism spectrum disorders in general education classes: Teachers were asked to first nominate students who best fitted a description of a person with an autism syndrome disorder, and then complete the Autism Syndrome Screening Questionnaire for every student in their classroom. Very high agreement was found between the two methods (more than 90%); however, the nomination method took 15 min per class on average, which is considerably less than the several hours teachers spent completing the screening questionnaires.

Nomination strategies and factors affecting the nomination results should be examined more thoroughly. The wording of the descriptions could be further modified to capture more than the small fraction of children with the most serious problems. As Dowdy, Doane, Eklund, and Dever (2013) demonstrated in their comparison of the teacher nomination method and universal screening, teacher nominations tend to identify less students with behavioral and emotional problems than universal screening does. They also pointed out that male gender, office discipline referrals, lower study habits and cooperation grades resulted in higher probability of the student being identified by either of the methods. Kroes, Veerman, and De Bruyn (2004) explored possible distortions of the child behavior evaluation stemming from the informant’s personality characteristics. The level of neuroticism of the teacher (as an informant) was positively associated with his/her ratings of child behavior problems.

7. CONCLUSION/DISCUSSION

To summarize the results obtained on the population-based sample and to compare them with previous findings, the overall degree of cross-informant agreement was low, particularly for ratings of boys. The highest degree of agreement both on problem behavior scales and on single items was found for parent-teacher pairs. The median correlation of 0.34 for the problem scales was even higher than the mean parent-teacher ratings correlation of 0.27 found by Achenbach et al. (1987) in their meta-analysis.

The overall agreement between teachers and children’s ratings was very low, with the exception of moderate agreement on the Externalizing Behavior scales. The median correlation
of 0.12 did not even reach the mean correlation of 0.20 for teacher-child pairs in the meta-analysis by Achenbach et al. (1987). Because teachers report fewer problems – and this is particularly the case for the internalizing problems – than children do, the low agreement between teachers and children may be due to a lack of variance in teachers’ ratings. Moreover, the children may be more likely to display internalizing behaviors in presence of their parents than in front of their teachers. This result is in accordance with previous findings (e.g. Stanger & Lewis, 1993; Achenbach et al., 1987) suggesting better agreement on externalizing behaviors than on internalizing behaviors, especially between teachers and other raters.

In spite of a limited validity of the Kappa coefficients due to low frequencies of teacher’s ratings in the nomination technique, we may assume that teachers were able to identify a certain proportion of children who rated themselves as having problems. This was particularly true for externalizing behaviors. This finding is consistent with our findings on the teacher-child agreement on the scales and items of the TRF and the YSR as described and discussed above. Our findings about the nomination technique are consistent with the conclusion of the above mentioned study for Dwyer et al. (2006), which found higher predictive power of a simple nomination question for externalizing problems score than for the internalizing problems and total problems scores. Correspondingly, Cunningham and Suldo (2014) found around 50% missing rates when identifying children with elevated levels of internalizing problems like depression and anxiety by the teacher nomination technique. As cost-efficient and timesaving the teacher nomination technique might be, for most problems it should not be used as the only screening method. The risk of misclassification can be lowered by complementing the data with peer nominations, self-reports, or parent reports.

Our findings emphasize the importance of obtaining information on children’s emotional and behavioral problems from multiple informants. The agreement between parents’, teachers’ and self-reports is too low to suggest that one source can substitute for another; the assessment of children must take account of variance across situations and informants on which the assessment depends.

REFERENCES


Cross-informant agreement and teacher nomination technique in the assessment of children behavior problems


ACKNOWLEDGEMENTS

The study was supported by Masaryk University (project MUNI/A/0887/2012).

AUTHOR(S) INFORMATION

Full name: Helena Klimusová
Institutional affiliation: Institute of Psychology, Faculty of Arts, Masaryk University
Institutional address: A. Nováka 1, 602 00 Brno, Czech Republic
Biographical sketch: Helena Klimusová is a senior lecturer in the Institute of Psychology at Masaryk University. She teaches methodology, statistics, general psychology and ethics in psychology. Her research interest include methodology and assessment methods, health-related behavior, and ethical issues of psychology.

Full name: Iva Burešová
Institutional affiliation: Institute of Psychology, Faculty of Arts, Masaryk University
Institutional address: A. Nováka 1, 602 00 Brno, Czech Republic
Biographical sketch: Iva Burešová is a senior lecturer in the Institute of Psychology at Masaryk University. She teaches developmental psychology, counselling psychology and work psychology. Her research interest include assessment methods, self-harm at children and adolescents, health-related behavior.
Full name: Ivo Čermák
Institutional affiliation: Institute of Psychology, The Academy of Sciences of the Czech Republic
Institutional address: Veverí 97, 602 00 Brno, Czech Republic
Biographical sketch: Professor Ivo Čermák is a senior scientist at the Institute of Psychology of the Academy of Sciences of the Czech Republic. He also teaches at the Faculty of Social Studies, Masaryk University. His research interest include qualitative methodology, narrative psychology, projective methods, developmental psychology, psychology of the art, aggression.