

## Preschoolers' Psychosocial Problems: In the Eyes of the Beholder? Adding Teacher Characteristics as Determinants of Discrepant Parent–Teacher Reports

Turid Suzanne Berg-Nielsen · Elisabet Solheim · Jay Belsky · Lars Wichstrom

Published online: 13 November 2011

© The Author(s) 2011. This article is published with open access at Springerlink.com

**Abstract** In this study, we explored informant characteristics as determinants of parent–teacher disagreement on preschoolers' psychosocial problems. Teacher characteristics were included in the analyses, in addition to child and parent factors. Psychosocial problems of 732 4-year olds from a Norwegian community sample were assessed by parents and teachers (CBCL-TRF). Furthermore, teachers reported on their education, experience and relationship to the child. Parental stress and psychopathology were also measured. Teachers rated children considerably lower than their parents did, especially on internalizing problems. When teachers rated more child problems, this was strongly associated with conflict in the teacher–child relationship, which predicted disagreement more than other factors. The highest agreement was on boys' externalizing problems. Girls' behavior was rated much lower by teachers than boys' behavior compared to parents' ratings. Possible teacher perception biases are discussed, such as teacher–child conflict, non-identification of internalizing problems, and same-gender child preference.

**Keywords** Informant · Discrepancy · Disagreement · Psychopathology · Child · Preschool

---

T. S. Berg-Nielsen (✉)

Regional Center for Child and Adolescent Mental Health, Faculty of Medicine, Norwegian University of Science and Technology, Postbox 8905, Trondheim, Norway  
e-mail: turid.suzanne.berg-nielsen@ntnu.no

T. S. Berg-Nielsen

RBUP, Faculty of Medicine, NTNU, Postbox 8905, MTF5, 7491 Trondheim, Norway

E. Solheim · L. Wichstrom

Department of Psychology, Norwegian University of Science and Technology, Trondheim, Norway

J. Belsky

Human and Community Development, University of California Davis, Davis, CA, USA

## Introduction

Understanding diverging informant reports of children's mental health problems becomes pertinent with preschoolers. Valid self-reports by preschool children are usually not available, and the identification of problems often relies entirely on reports from parents and teachers/caregivers. Informants' endorsements of diagnoses can have far-reaching consequences for the identification and referral of troubled children [1]. Discrepancies between informants, however, are also the rule with older children and remain one of the most poorly understood phenomena in child diagnostic research [2, 3]. The first meta-analysis of parent–teacher agreement with older children showed a correlation of 0.27 between the Child Behavior Checklist (CBCL) and the Teacher Report Form (TRF) [4]. Later studies showed similar results [5, 6] or no concordance whatsoever between parents and teachers [7].

So far, one cannot assert that one observer-informant is closer than the other to a gold standard of estimating a child's mental health [1, 8]. Nevertheless, parents' ratings have shown to be better predictors than teachers' of future child diagnoses [9]. Disagreement between observer-informants as such does not represent a risk factor for a later adverse child outcome [9]. On the contrary, agreement on high ratings of CBCL and TRF total problems scores is a strong predictor of a child's later need for professional help [10, 11]. Viewed against this background, agreement on a child's total problem level could be regarded as a measure that brings us closer to the abovementioned gold standard. Consequently, it is important to identify factors that might weaken that agreement.

Achenbach et al. [4] argued long ago that an observer-informant discrepancy does not necessarily represent an actual disagreement because it could derive from inconsistency of child behavior across settings, often called situational specificity, with children behaving differently at home and at school. This has been the predominant explanation of informant disagreement and is partly supported by evidence indicating that cross-informant agreement increases when informants interact with the child in the same situations (e.g., between mothers and fathers) [12]. To test the validity of context-specific child behavior as an explanation of parent–teacher disagreement, however, one would need numerous informant-observers who could rate child behavior over time in naturalistic settings both at home and at school, without being associated with either one. Such a study would hardly be viable, given its unreasonable premise that an observing stranger would nearly have to reside within the private sphere of a family. A different approach would be to investigate how much of the variance in informant disagreement could be accounted for by observer characteristics. Such characteristics, which would possibly represent reporting biases, could explain disagreement that is not accounted for by situational specificity of child behavior.

Despite the low agreement among informers, informant-based methods of preschool behavioral assessment are wide-spread and suggest a pressing need for research that examines informant influences on ratings [13]. Konold and Pianta [14] addressed this issue with a Latent Profile Analysis of CBCL and TRF ratings of first graders from the NICHD study and found that ratings were heavily influenced by informants.

### Determinants of Parent–Teacher Disagreement

In previous CBCL-TRF discrepancy research, the focus has mainly been on parent characteristics as determinants of informant discrepancies and less on child factors. Teacher characteristics as determinants of CBCL-TRF disagreement on children's problems have

been addressed in one study [15], which showed that teachers' familiarity and contact with the children enhanced parent–teacher agreement.

Each of the following three sets of explanatory factors merits some attention.

### *Child Characteristics*

Child factors that may influence informants' agreement, other than context-specific behavior, are type and severity of problems, age and gender. Child gender however, has not been reported as affecting informants' agreement. With older children, parents and teachers agree more on externalizing than internalizing disorders, as might be expected because observable behavior is easier to detect than internal states [5, 16–18]. Although numerous studies provide data on cross-informant agreement of the CBCL and TRF, the authors have not found studies that specifically address the determinants of disagreement in a large community sample of preschoolers.

### *Parent Characteristics*

The depression–distortion hypothesis [19] has been confirmed as a determinant of informant disagreement in several studies of clinical samples; high levels of maternal depressive symptoms are related to over-reporting of children's behavior problems compared to ratings by the children or their teachers [3, 20, 21]. However, parents' possible reporting biases may vary as a function of sample selection. Due to a desire to elicit services, parents drawn from clinical samples may over-report problems [1, 22]. In contrast, parents drawn from a community sample may under-report to avoid stigmatizing their child or attracting unwanted involvement from professionals. Among other parent factors, stress is also related to over-reporting of externalizing problems compared to other informants [1, 5]. In addition to addressing this issue in the present study, deviant personality traits of parents are also examined as possible determinants of parent–teacher disagreement, given the potential role of projection bias in parental ratings [23].

### *Teacher Characteristics*

Although teachers usually have far more experience than parents, due to the observation of numerous children, experience does not preclude that teachers, like parents, may be prone to reporting biases. However, only one study has examined the possibility that teacher characteristics may be the drivers of parent–teacher disagreement [15]. Because a teacher's education, the length of time they have known a child, prior experience with children, the kind of relationship they have with a child and gender issues (e.g., same-gender preference) might all be factors that influence teacher reports, these are also included in the research reported herein to estimate which determinants are more predictive of disagreement about preschoolers' mental health problems.

### Discrepancy Measurement Issues

De Los Reyes and Kazdin [1] urged researchers to pay careful attention to the measure of informant disagreement chosen. They tested the three most common ways of measuring discrepancies between informers. In the *raw difference score*, one informant's raw score is subtracted from the other informant's raw score. In the *standardized difference score*, each

informant's rating is converted into a  $z$  score by placing it on a distribution relative to the rest of the same informant's ratings in the sample. The informant's converted rating is then subtracted from the other informant's converted rating. The *standardized difference score* is often used to place both informants' ratings on the same metric, or scale of variability (e.g., the  $z$  distribution), because doing so enhances the interpretability of the score. With the *residual difference score*, one informant's rating is used as an independent variable in a regression model to predict the other informant's rating as a dependent variable. The difference between the ratings predicted by the independent variable and the dependent variable informant's actual rating becomes the *residual difference score*, which is then standardized.

De Los Reyes and Kazdin [1] showed how the mathematical properties of the three measures of informant discrepancies can lead to different conclusions, depending on the differences between variances of informants' ratings or the correlation between them. E.g., when variances of groups of informants are equal, *raw difference scores* and *standardized difference scores* will bring forth the same results when examining informant discrepancies with other variables. However, when the variances of one group of informants are much larger than the variances of the other group of informants, the *raw difference scores* and the *standardized difference scores* will yield maximally different findings when discrepancies are related to other variables of interest. Furthermore, the results of relating the *residual difference scores* with other variables are dependent on the correlation between informants' ratings. Of the three discrepancy measures De Los Reyes and Kazdin [1] compared in their study, standardized difference scores produced the most consistent estimates among informant discrepancies and informant characteristics. Furthermore, it was the only discrepancy measure that correlated equally with each of the informants' ratings from which it was created, in concordance with the assumption that no informant more than the other represents a "gold standard" of estimating child mental health. Hence, De Los Reyes and Kazdin [1] recommended that future investigations use the standardized difference score when discrepancies are related to informant characteristics.

## Research Questions

This study's main objective was to explore determinants of parent–teacher (dis)agreement on preschoolers' mental health problems in a community sample by including relevant teacher characteristics in addition to child and parent factors. The teacher characteristics in this study were education, experience with children in general, length of acquaintance with the target child, and perception of the teacher–child relationship. The child characteristics included age, gender and problem severity. Furthermore, potential predictors of parent–teacher disagreement included well-documented parent characteristics, such as stress and psychopathology, in addition to their own experience with children.

An additional research question was addressed because it is especially relevant for identification and referral practices: which children did teachers and parents mainly disagree on, the ones whom they perceived to be troubled or those who appeared well-adjusted?

## Method

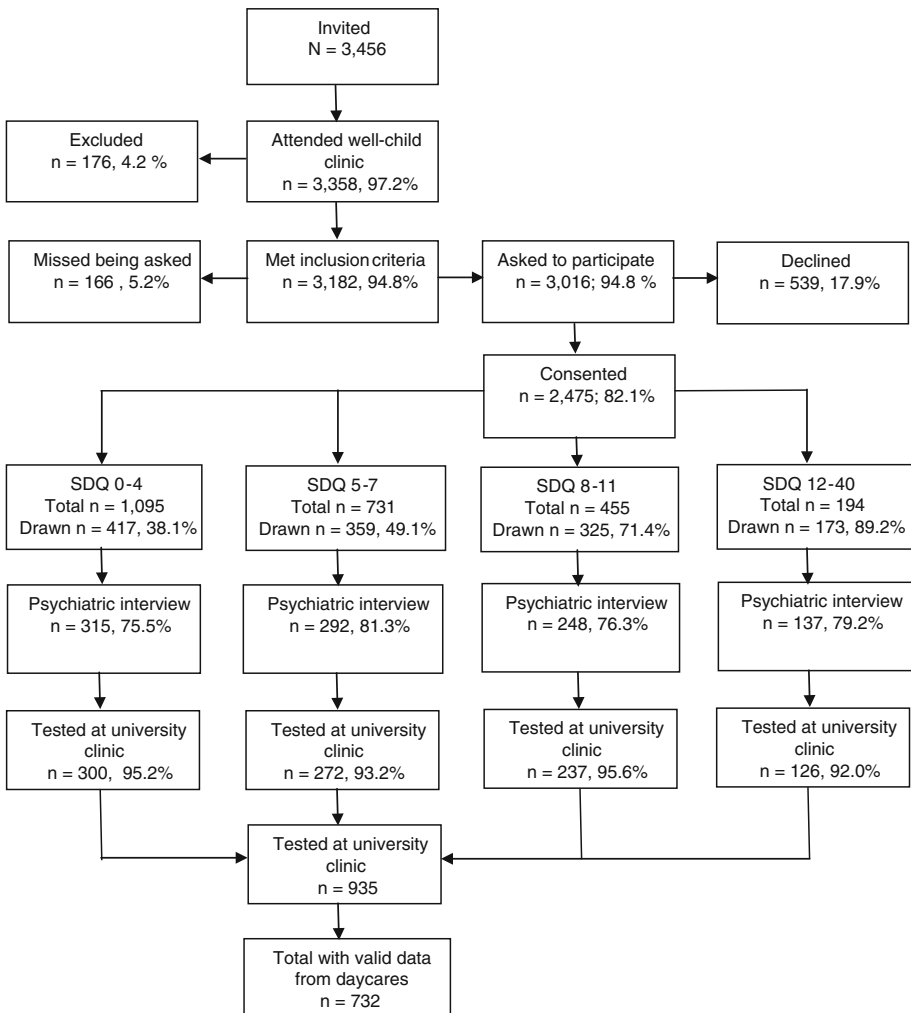
### Participants and Recruitment

All children born in 2003 or 2004 and their parents living in the city of Trondheim, Norway were invited to participate in the study ( $N = 3,456$ ). A letter of invitation together

with the Strengths and Difficulties Questionnaire (SDQ) [24, 25] was sent to their homes. The parents brought in the completed SDQ when attending their scheduled appointment for the ordinary community health checkup for 4-year olds. A flow-chart describing the recruitment procedure and the participation rates is depicted in Fig. 1.

As can be seen, almost everyone who was eligible for the study appeared at the city’s well-child clinics, meaning that the sample is, in practice, a community sample. Parents with insufficient proficiency in Norwegian to fill out the SDQ screen were excluded. The health nurse at the well-child clinic informed the parent about the study using procedures which were approved by the Regional Committee for Medical and Health Research Ethics, and obtained written consent to participate (5.2% of eligible parents were missed).

SDQ Total Difficulties scores (20 items) were divided into four strata (cut offs: 0–4, 5–7, 8–11, 12–40). Using a random number generator, defined proportions of parents in



**Fig. 1** Flowchart of sample recruitment

each stratum were drawn to participate in a structured diagnostic interview concerning the child's mental health. The drawing probabilities increased with increasing SDQ scores, i.e., they were 0.37, 0.48, 0.70, and 0.89 in the four strata, respectively. Of the 1,274 parents who consented to participate at the well-child clinics, 992 (79.5%) showed up for the subsequent assessment at the university clinic. This drop-out rate was unchanged across the four SDQ-strata (Chi-sq. = 5.70,  $df = 3$ , NS) or gender (Chi-sq. = 0.23,  $df = 1$ , NS). The sample, adjusted for stratification, was compared with register information from Statistics Norway [26] on all parents of 4-year olds in Trondheim in the years 2007 and 2008. The only difference was that our sample contained significantly more divorced parents (6.8%) than the Trondheim population (2.1%). In our sample, the educational level, which highly correlates with SES, was virtually identical to the population's level. Furthermore, the population of Trondheim is similar to the national average on several key indicators: the average gross income per person is 99.5% of the national average, the employment rate is identical to the national rate, and 80.0% of the households are two-parent families compared to a national average of 81.4% [26].

Descriptive information about the sample is shown in Table 1.

A subsequent comprehensive assessment took place within a couple of weeks following the visit at the well-child clinics. This assessment included questionnaires for parents and interviews and observation measures of the child. The research assistants that conducted the assessment ( $n = 7$ ) had at least a bachelor's degree in relevant fields and extensive prior experience in working with children and families. To be included in the current study, the children's preschool teachers had to have returned the completed Student-Teacher Relationship Scale (STRS) [27, 28] and the TRF of the ASEBA for preschoolers [29]. 10% of the preschool teachers did not return the forms and an additional 11% had incomplete data so the final sample for this study included 732 children, with 359 girls and 373 boys who had a mean age of 54.7 months ( $SD = 3.02$ ). This sample did not differ from the 1,274 children drawn to participate in the study on the SDQ ( $t = 0.613$ ,  $p = .54$ ), nor did they differ from the 992 parents interviewed in terms of child age ( $t = -1.335$ ,  $p = .183$ ) or gender ( $\beta = -0.07$  (0.12),  $p = .929$ ), family socioeconomic status ( $t = 0.401$ ,  $p = .681$ ), parent education ( $t = -1.484$ ,  $p = .138$ ), family income ( $t = 1.528$ ,  $p = .127$ ), parent's ethnicity (mother:  $\beta = 0.47$  (0.95),  $p = .838$ ; father:  $\beta = 0.07$  (0.24),  $p = .761$ ) or marital status ( $\beta = -0.03$  (0.05),  $p = .594$ ).

A total of 95% of the children were in center-based daycare (see Table 1), at which the person who knew the child best was asked to complete the TRF. Among these, 83.7% were educated preschool teachers, while 16.3% had other educations. Preschool teachers in Norway have a separate 4-year educational program that leads to a bachelor's degree. A total of 86.2% were women and 13.5% men, and they had a mean age of 38.2 years, which ranged from 22 to 70 ( $SD = 8.7$ ). Overall, they had extensive experience working with children (mean = 13.3 years,  $SD = 8.6$ ) and had known the child in question an average of 18.8 ( $SD = 10.95$ ) months. Norwegian daycare centers have only two main "classes" or departments, one for those younger than 3 years and another for children 3–5 years. Each child is assigned a primary contact among the adults, with the aim to keep the relationship consistent. Thus, there is greater stability in the child-caregiver relationship than in countries where children change preschool classes 1 or 2 times a year.

In appreciation of their participation, families were given a lottery ticket with the prize being a free trip for the family (their choice) worth \$7,000. During the examination children were given snacks and small prizes (e.g., blowing bubbles) as encouragement.

The study was approved by the Norwegian Regional Committee for Medical and Health Research Ethics.

**Table 1** Sample characteristics

Characteristic		%
Gender of child	Male	49.1
	Female	50.9
Gender of parent informant	Male	15.2
	Female	84.8
Ethnic origin of biological mother	Norwegian	93.0
	Western Countries	2.7
	Other Countries	4.3
Ethnic origin of biological father	Norwegian	91.0
	Western Countries	5.8
	Other Countries	3.2
Childcare	Official daycare center	95.0
	Other	5.0
Biological parents' marital status	Married	56.3
	Cohabiting >6 months	32.6
	Separated	1.7
	Divorced	6.8
	Widowed	0.2
	Cohabiting <6 months	1.1
	Never lived together	1.3
Inform. parent's socio-econ. status	Leader	5.7
	Professional, higher level	25.7
	Professional, lower level	39.0
	Formally skilled worker	26.0
	Farmer/fisherman	0.5
Parent's completed education	Unskilled worker	3.1
	Not completed junior high school	0
	Junior high school (10th grade)	0.6
	Some educ. after junior high school	6.1
	Senior high school (13th grade)	17.3
	Some educ. after senior high school	3.4
	Some college or university education	7.6
	Bachelor degree	6.2
	College degree (3–4 years study)	33.6
	Master degree or similar	20.3
Ph.D completed or ongoing	4.4	
Households' gross annual income	0–225' NOK (0–40' USD)	3.3
	225'–525' NOK (40'–94' USD)	18.4
	525'–900' NOK (94'–161' USD)	51.6
	900' + NOK (161' + USD)	26.7

### Child Measures

Developed in Great Britain [24] and extensively evaluated in many countries and cultures [30], the SDQ 4-16 version is a 31-item measure of psychiatric symptoms. It has been proven to function well as a screening device for evaluating children's mental health [31].

In this study, the Total Difficulties score comprising 20 items was used for screening [32]. The SDQ was translated into Norwegian in 2001 [33], and the Norwegian version has been validated in several large studies [34, 35]. In our study, Cronbach's alpha for the Total Difficulties score was 0.77.

The Norwegian version of The Child Behavior Checklist (CBCL/1.5–5) was used in the study. CBCL is a standardized form that assesses parent-reported behaviors, problems and competencies in children 1½–5 years [29]. Although most research with CBCL and TRF have focused on school children and adolescents, there are more than 200 studies of the preschool forms [36]. The psychometric qualities of the CBCL/1.5–5 (and C-TRF/1.5–5) have been confirmed with test–retest reliability in the 0.80s and 0.90s [37–39]. There are seven syndromes that constitute the Total Problems Scale [40], which seems to give a good indication of psychosocial problems in preschoolers [41]. However, the preschool version has not yet been validated in a larger Norwegian sample. The clinical scales of the CBCL contain a Total Problems score, two broadband dimensions, internalizing problems and externalizing problems, and seven other scales. Cronbach's alpha in our sample for CBCL internalizing problems was 0.83, CBCL externalizing problems = 0.89, and CBCL total problems = 0.93.

The Norwegian version of the Caregiver–Teacher Report Form for ages 1.5–5 (C-TRF) covers the same dimensions as CBCL (except for Sleep Problems) [29]. The reliability of the C-TRF has proven to be good-to-sufficient in large German and Chinese samples [42, 43]. Cronbach's alpha in our sample was for TRF internalizing = 0.86, TRF externalizing = 0.95, and TRF total = 0.95.

### Teacher Measures

The STRS [27, 28] was used to assess the teacher's perception of the quality of his or her relationship with a given child. The STRS has been translated into Norwegian in accordance with standard translation protocol [44, 45]. The STRS comprises 28 items that are distributed across three subscales: Conflict, Closeness and Dependency. In our sample Cronbach's alpha ranged from 0.81 to 0.87. Its validity has been documented in previous studies [46, 47]. In the current study, we measured teacher–child conflict using a slightly modified, adapted version of the STRS [47]. Information on teacher education and experience with children was obtained in a separate questionnaire. Not including our CFA of the STRS with this sample [47], three earlier studies have used the full scale STRS with preschool samples (under age five) [48–50].

### Parent Measures

Parental stress was measured with the Parenting Stress Index (PSI) [51], a 120-item scale of parents' perceived stress (in our sample, alpha = 0.85) that yields scores on two subscales. The Norwegian version was translated by Ronning and Abidin. In the Nordic countries there is only one validation study, from Sweden [52]. One of the two subscales is related to the child (i.e., the child domain; e.g., child demandingness), and the other concerns stress experienced in the parent role and parental functioning (i.e., the parent domain).

Parental symptoms of depression were measured with the self-reported Beck Depression Inventory-II (BDI) [53] (alpha = 0.87 in our sample). The BDI-II consists of 21 items encompassing nine symptoms of depression corresponding to the DSM-IV criteria. A meta-analysis of BDI has been performed and the inventory shows good reliability and



high convergent and discriminant validity [54]. The BDI was first translated to Icelandic before further translation to the other Nordic languages. Iceland is the only Nordic country that has a validation study of the BDI [55].

Parental symptoms of anxiety were measured with the Beck Anxiety Inventory (BAI) [56] Cronbach's alpha in our sample was 0.82. The BAI has been translated into Norwegian with the approval of its distributors. It also consists of 21 items measuring two factors: physiological and cognitive symptoms of anxiety. It has shown good convergent and discriminant validity in clinical and non-clinical samples in several countries. There is no known Norwegian validation study.

Parents' deviant personality traits were measured with The DSM-IV and the ICD-10 Personality Questionnaire (DIP-Q) [57]. The DIP-Q is a 140-item true/false self-report questionnaire designed to measure all ten DSM-IV and all nine ICD-10 disorders. In the DSM-IV, the personality disorders are divided into three clusters: Cluster A, Odd/Eccentric (Theta = 0.95), Cluster B, Dramatic/Erratic (Theta = 0.91), and Cluster C, Anxious (Theta = 0.91). Since Cronbach's alpha assumes that data are continuous, the Theta coefficient was used to measure reliability. When Cronbach's alpha is measured on categorical scales with less than six categories in addition to a skewed distribution of scores, alpha underestimates the internal consistency [58]. A reliability test that takes into account that data are ordered categorically, such as the Theta, and not continuously more accurately portrays the reliability and is therefore recommended [59]. In this study, we used DIP-Q cluster scores as possible determinants of parent–teacher disagreement. The validation of the DIP-Q cluster scores for screening purposes has been documented, and DIP-Q shows acceptable test–retest reliability [60, 61].

## Statistical Analyses

In this study, we used the standardized difference score to measure informant discrepancies: the teachers'  $z$  score was subtracted from the parents'  $z$  score. To determine whether parents reported more problems than teachers or vice versa, all analyses were repeated, using the difference between informants' unstandardized ratings. Positive scores indicated that parents reported more problems than teachers, while negative scores indicated that teachers reported more problems than parents.

Because we had a stratified sample, all results were weighted back to represent true population estimates using a Huber-White sandwich estimator. A plan file was created in PASW Statistics 18 Complex Samples (CS), with the samples in the four strata weighted and used in all analyses. Parameter estimates of  $t$ -statistics of mean difference of discrepancies between parents and teachers on girls and boys and internalizing and externalizing problems were used. The PASW CS General Linear Model was used with parameter estimates of the separate effect of each predictor and Wald  $F$  statistics for the effects of child, parent, and teacher characteristics on the model. In consideration of the sample size and the number of predictors involved, level of significance was set at  $p < .01$ .

To obtain an indication of the direction of the discrepancies (i.e., whether the discrepancies increased or decreased as a result of parents or teachers reporting more problems (or less), each independent variable that proved significant in the model was recoded into 4 categories along the 25th, 50th and 75th percentiles. CBCL and TRF means were calculated across the four groups of low-to-high scores of each independent variable that was significant in the model, rendering it possible to compare whether parents or teachers reported more with increasing values of the independent variable.

To test whether parent–teacher disagreement differed depending on severity of child problems, i.e., whether parents and teachers disagreed more on troubled than well-adjusted children, the TRF Total Problem Scale (TPS) was first split into two groups along the median = 8. We chose the TRF scale for this analysis and not the CBCL as teachers rated the children lower than parents on all scales, thus rendering a more cautious estimate of psychopathology. The “low” group with scores 0–8 comprised 47% of the whole sample. Of those children who had TPS scores above 8, 10% scored higher than 33, which was chosen as the high scoring group since 33 is a rough estimate of clinical cut-off [29]. Scores from 9–32 constituted the “middle” group. Crosstabs were then calculated to obtain percentages of agreement across low, medium and high CBCL and TRF scores.

## Results

The overall parameter estimate of correlations between the CBCL and TRF total problems in the whole sample was 0.26, which corresponds almost perfectly with the 0.27 figure originally reported by Achenbach for older children [4]. Thus, parent–teacher disagreement does not appear to increase with younger children, at least according to this between-study comparison. As can be seen in Table 2, parents consistently reported more child problems than teachers. Teachers were especially low on ratings of internalizing disorders for both boys and girls compared to parents.

### Disagreements on Internalizing and Externalizing Problems

The parent–teacher disagreement was significantly larger for internalizing problems than externalizing problems due to teachers’ low ratings of internalizing problems (Table 3).

Regarding externalizing problems, more disagreement existed in the case of girls, with teachers rating girls’ behavior lower than parents. Parents and teachers were more in agreement on boys’ externalizing problems. Parent–teacher disagreement on Total problems followed the same pattern with significantly more disagreement about girls than boys.

### Determinants of Disagreement

40.4% of the variance in the parent–teacher discrepancy on total child problems was explained by the General Linear Model in Table 4.

A teacher-perceived conflict with a child contributed more than other predictors to the discrepancy score (Wald  $F = 44.247$ ). The conflict effect was present not only for externalizing problems but also for internalizing (and thus total) problems.

The negative parameter estimate of the teacher–child conflict shows that the discrepancy decreases. In addition to the teacher–child conflict, parenting stress (the child domain) was also related to the discrepancy score. It showed a positive parameter, which indicates that the CBCL-TRF discrepancy increased.

To determine the direction of the discrepancies, i.e., whether parents or teachers reported more or less child problems than the other, CBCL-TRF means were calculated across four levels of scores (cut points along the 25th, 50th, and 75th percentiles) of the two significant independent variables in the model. Thus, the direction of the CBCL-TRF scores became apparent, whether CBCL or TRF scores went up or down depending on the level of teacher–child conflict or parental stress (Table 5).

**Table 2** CBCL and TRF mean raw scores (SD) by parents, fathers, mothers and teachers for all children, girls and boys. T-statistics of difference between how boys and girls were rated and between mothers' and fathers' scores

	All	Girls	Boys	t-stat.	
				Girls/boys	Mothers/fathers
Parents (CBCL)					
Total problems	17.049 (0.33)	16.600 (0.46)	17.478 (0.51)	-1.105	-
Internalizing	4.944 (0.12)	5.008 (0.18)	4.944 (0.19)	0.439	-
Externalizing	6.156 (0.14)	5.752 (0.21)	6.472 (0.21)	-2.339*	-
Mothers (CBCL)					
Total problems	17.453 (0.49)	17.1512 (0.70)	17.5122 (0.71)	-1.091	-
Internalizing	5.121 (0.19)	5.1293 (0.26)	5.0213 (0.26)	0.276	-
Externalizing	6.321 (0.21)	5.8574 (0.30)	6.5569 (0.30)	-2.375*	-
Fathers (CBCL)					
Total problems	16.294 (1.21)	14.3872 (1.32)	17.2894 (1.86)	-1.186	1.412
Internalizing	4.520 (0.47)	4.1401 (0.44)	4.7022 (0.75)	-0.618	1.643
Externalizing	5.687 (0.52)	4.9562 (0.65)	6.0221 (0.75)	-1.195	1.404
Teachers (TRF)					
Total problems	13.769 (0.42)	11.326 (0.53)	15.875 (0.67)	-5.350***	-
Internalizing	1.216 (0.06)	1.237 (0.09)	1.236 (0.08)	0.470	-
Externalizing	4.595 (0.17)	3.430 (0.20)	5.509 (0.27)	-6.104***	-

\*  $p < .05$ ; \*\*\*  $p < .001$ **Table 3** CBCL minus TRF (and standardized difference) and t-values of differences in mean discrepancies between externalizing and internalizing disorders and child gender differences

	All	Girls	Boys	Gender difference t
CBCL minus TRF				
Total problems	3.351 (-0.078)	5.272 (0.037)	1.348 (-0.198)	3.363***
Internalizing	3.699 (-0.073)	3.799 (-0.055)	3.673 (-0.092)	0.585
Externalizing	1.592 (-0.082)	2.257 (0.024)	0.909 (-0.191)	3.089**

Difference in CBCL-TRF discrepancy scores for all on internalizing and externalizing problems:  $t = 18.674***$ \*\*  $p < .01$ ; \*\*\*  $p < .001$ 

The decreased discrepancy between CBCL-TRF when teachers experienced conflict with a child was due to teachers reporting more child problems as can be seen in Table 5. Parents also reported somewhat increased problems when there was a high teacher child-conflict; however, teachers reported considerably more. Although teachers, overall, reported fewer child problems than parents (Tables 2, 3), when they did experience high conflict with a child, they reported even more problems than parents did (Table 5).

Parents reported more problems than teachers when they experienced a high stress level, which contributed to a larger informant discrepancy. Teachers also reported more child problems when parents were more stressed, but not enough to diminish the gap to the high problem scores of the parents.

**Table 4** GLM with parameter estimates PE and (SE) and Wald F of child (C), parent (P) and teacher (T) characteristics on standardized discrepancies of CBCL-TRF total, internalizing, and externalizing problems

Z discrep. of	Total problems		Internalizing		Externalizing	
	PE (SE)	Wald F	PE (SE)	Wald F	PE (SE)	Wald F sig
Model R <sup>2</sup>	0.404		0.264		0.374	
C gender	-0.023 (0.140)	0.028	-0.291 (0.160)	3.282	0.122 (0.143)	0.728
P depression	0.014 (0.015)	0.876	0.013 (0.016)	0.619	0.015 (0.015)	0.990
P anxiety	-0.013 (0.033)	0.163	0.019 (0.038)	0.248	-0.028 (0.034)	0.679
P Cluster A	0.000 (0.036)	0.000	0.024 (0.041)	0.339	-0.021 (0.041)	0.271
P Cluster B	-0.013 (0.029)	0.198	0.013 (0.031)	0.169	-0.050 (0.029)	2.923
P Cluster C	0.062 (0.030)	4.209	0.046 (0.037)	1.525	0.066 (0.030)	4.937
P stress 1 <sup>a</sup>	0.022 (0.006)	15.176***	0.012 (0.006)	3.447	0.026 (0.006)	19.116***
P stress 2 <sup>b</sup>	0.001 (0.005)	0.062	0.003 (0.005)	0.494	0.003 (0.005)	0.358
P experience	0.024 (0.086)	0.076	0.036 (0.091)	0.162	0.144 (0.083)	3.004
T experience	0.000 (0.001)	0.018	0.000 (0.001)	0.357	0.001 (0.001)	1.570
T knowledge <sup>c</sup>	-0.005 (0.005)	0.000	0.001 (0.006)	0.011	-0.002 (0.006)	0.088
T education	0.045 (0.083)	0.292	-0.022 (0.120)	0.034	0.042 (0.072)	0.342
T-C conflict	-0.163 (0.024)	44.247***	-0.119 (0.023)	27.661***	-0.136 (0.024)	32.596***
T-C closeness	0.034 (0.017)	4.019	0.004 (0.019)	0.054	0.003 (0.015)	4.880
T-C depend	-0.006 (0.046)	0.018	-0.044 (0.045)	0.961	0.010 (0.046)	0.048

<sup>a</sup> Parenting stress—child domain, <sup>b</sup> parenting stress—parent domain, <sup>c</sup> how long the teacher had known the child  
 \*\*\*  $p < .001$

### Parent–Teacher Disagreement and Child Problem Severity

To test whether child problem severity was associated with the degree of parent–teacher disagreement, the CBCL/TRF total problem scales were divided into three groups: low, medium and high scores. These were then placed in crosstabs in Table 6. When teachers rated children low, 32% of the parents agreed and also rated the children low; however, 58.9% of the parents scored their children to be in the medium range of problem severity while teachers perceived no or few problems. When teachers rated the children high, only 24.9% of the parents agreed.

As can be seen in the right half of Table 6, teachers agreed somewhat more with parents' ratings than the other way around. When parents rated the children low, more than half of the teachers agreed (54.5%). However, when parents rated the children with high problem severity, only 22.2% of teachers agreed, and as many as 39% of the teachers perceived no or few problems.

## Discussion

### Teacher-Perceived Conflict as a Determinant of Disagreement

Overall, teachers in our sample reported considerably fewer problems in children than did parents. However, when they did report more problems, teacher-reported conflict in the

**Table 5** CBCL & TRF means for total, internal, and external problems and t-statistics of CBCL-TRF differences across 4 groups (from low to high scores) of the 2 significant ( $p < .01$ ), independent variables in Table 4, teacher–child conflict and parental stress

	Low	Low–medium	Medium–high	High
STRS teacher–child conflict				
CBCL total problems	15.215	17.276	17.016	19.636
TRF total problems	7.049	6.891	13.793	31.014
t-CBCL/TRF total differ.	12.604***	13.079***	4.806***	–10.942***
CBCL internalizing	4.417	4.532	5.490	5.430
TRF internalizing	0.826	0.713	1.253	2.440
t-CBCL/TRF intern. differ.	15.310***	14.660***	14.031***	7.826***
CBCL externalizing	5.021	5.952	5.926	8.507
TRF externalizing	2.023	2.029	4.802	10.984
t-values CBCL/TRF extern.	11.354***	12.123***	3.457**	–5.970***
PSI parental stress (child domain)				
CBCL total problems	8.436	15.314	18.790	29.102
TRF total problems	10.426	14.122	15.926	18.438
t-CBCL/TRF total differ.	–6.432***	0.541	3.755***	11.819***
CBCL internalizing	2.326	4.359	5.140	8.222
TRF internalizing	0.938	1.011	1.328	1.590
t-CBCL/TRF intern. differ.	7.607***	14.961***	14.694***	19.466***
CBCL externalizing	2.279	4.944	6.639	10.799
TRF externalizing	3.464	4.261	4.895	5.522
t-CBCL/TRF extern. differ.	–7.414***	2.536*	5.987***	15.435***

\*  $p < .05$ ; \*\*  $p < .01$ ; \*\*\*  $p < .001$ 

teacher–child relationship was the main explanatory factor. One might assume that teacher-reported conflict especially affected parent–teacher disagreement when children had externalizing and defiant behavior, which causes a strain on any relationship. Conflict, however, was also the only significant predictor of disagreement on internalizing problems (Table 4), which resulted in teachers reporting more internalizing problems. Teacher–child conflict alone explained 26.4% of the variance in disagreement on internalizing problems. Thus, when teachers, compared to parents, perceived high conflict with a child, they reported more externalizing and internalizing problems (Table 5). Subsequently, the question arises of whether the teacher–child relationship represents a perception bias of a child’s problem level that reflects teachers’ child preferences. Such a bias might also be due to an attributional process; teachers’ perceived conflict with a child is attributed as a negative child characteristic. Yet, teachers, as parents, are obviously more often in conflict with children who show problem behavior, as was shown in a study that perceptions of those children’s problem behaviors explained over half of the variance in teachers’ reports of conflict with preschoolers [62]. However, the direction of effect is not the issue in this study. What is noteworthy is that teacher–child conflict increases the *discrepancy* between teachers’ and parents’ reports. Overall, parents usually report more child problems than teachers do, but if there is high teacher–child conflict, then teachers surpass even the parents in perceiving child problems, as evident in Table 5. Possibly strengthening a teacher-bias hypothesis is that teachers’ depression and feelings of low self-efficacy have been shown to influence their reports of conflict in the teacher–child relationship more than

**Table 6** Columns 1, 2, 3 of the table show percentage of parents who agreed with teachers' ratings along 3 scale groups of low, medium, and high CBCL-TRF scores of total, internalizing, and externalizing problems. Columns 4, 5, 6 of the table show percentage of teachers who agreed with parents' ratings

CBCL total problems (%)			TRF total problems (%)			
	Low	Medium	High	Low	Medium	High
TRF total				CBCL total		
Low	32.0	58.9	9.1	54.5	39.3	6.2
Medium	27.3	61.9	10.8	47.8	42.4	9.8
High	17.4	57.7	24.9	39.0	38.8	22.2
CBCL internalizing (%)			TRF internalizing (%)			
TRF intern.				CBCL intern.		
Low	16.4	24.3	59.3	77.5	19.2	3.3
Medium	12.8	29.0	58.3	65.2	24.7	10.2
High	3.2	17.0	79.8	62.9	19.3	18.6
CBCL externalizing (%)			TRF externalizing (%)			
TRF extern.				CBCL extern.		
Low	38.5	52.8	8.7	64.8	30.3	5.0
Medium	24.6	60.0	15.3	49.3	40.9	9.8
High	15.3	54.3	30.5	33.7	43.4	22.9

their experience and education [62]. In our study, none of the other teacher characteristics emerged as significant predictors of discrepancies, such as education, experience and how long they have known a child.

#### Teachers' Low Ratings of Internalizing Problems

The CBCL mean level of problems scores of the 4-year olds in this study are consistent with what has recently been found in a population-based Danish preschool sample [63]. The Norwegian teachers, however, rated children much lower on internalizing problems than their Danish counterparts. A benevolent interpretation is that Norwegian teachers are more tolerant and less worried about such problems in young children than parents and that they might be reluctant to declare internalizing behavior as problematic at this age. Sadness, tears and fears may simply be regarded as common and age-appropriate emotions for 4-year-olds in Norwegian daycare settings. Generally, most studies actually show that parent–teacher agreement is lower for internalizing problems than for externalizing [5, 15, 16, 64]. The common explanation for this has so far been that children more likely confide in their parents than their teachers about their emotional problems.

To argue that parent–teacher discrepancy on internalizing problems is due to situational specificity of child behavior is somewhat implausible. In the Danish study [63], parent–teacher agreement on internalizing problems with preschoolers was higher, although they did not specifically address informant discrepancies. Given the cultural similarities between Denmark and Norway, it is not likely that Norwegian preschoolers' internalizing behavior should be more context-specific than that of Danish children.

Teachers' low ratings on internalizing problems might also reflect a problem of detection; teachers may not be sufficiently attentive to or aware of young children's internalizing problems, failing to see them in the course of a busy day, even though symptoms of depression and anxiety manifest themselves in much the same way in preschoolers as in older children [65, 66].

#### Do female Teachers Have a Same-Gender Preference?

Teachers and parents disagreed more on the externalizing problems of girls than those of boys, with teachers rating the girls significantly lower than the parents (Tables 2, 3). We suggest two possible explanations that are not mutually exclusive. Girls' externalizing behavior may be more context-specific than boys' in that girls may show more externalizing behavior at home than in preschool compared to what might be characteristic of boys' behavior. This hypothesis may be due to more socially desirable behavior in girls than boys; girls maybe strive more to maintain socially acceptable behavior at school, but relax and act out as soon as they get home. The other explanation could be a possible same-gender bias among teachers. Of the raters from the daycares, 86.2% were female. An imminent question is whether their lower ratings of girls', but not boys', externalizing problems compared to parents' is due to more tolerance for the way girls' externalizing behavior manifests itself while boys' misbehavior is perceived as more bothersome. A possible same-gender favorable bias is only slightly evident with mothers; mothers' ratings of girls' behavior problems were higher than fathers' and differed even more from teachers' ratings, yet was somewhat lower than their ratings of boys (Table 2).

With boys' increased vulnerability for externalizing problems taken into account, a possible same-gender bias among preschool teachers gives cause for some concern, especially if it also implies that teachers actually favor girls' behavior to the disadvantage of boys. The issue could be addressed in future research, preferably including girls' and boys' own reports on symptoms, wellbeing in daycare and perceptions of their teachers (e.g., with the Berkeley Puppet Interview) which may add perspectives, and validate or weaken the hypothesis of teachers' same-gender preference.

#### Are Teachers' Reports more Valid Because of Extended Experience with Children?

Developmental researchers are inclined to regard teacher reports as more valid than parents' because teachers' experience with numerous children is presumed to afford them a better sense of age-appropriate child behavior. Furthermore, some researchers contend that symptoms and dysfunction are more likely to manifest themselves in the more demanding environment of school than at home [6]. That might explain why in one study, teachers of older children noticed internalizing problems more than parents [67] and also why teacher ratings can be more similar to the child's own ratings [5]. However, some children high on social desirability behave well in a school setting and instead act out at home.

Despite well-known parental reporting biases such as stress and depression [3, 19–21], clinicians still tend to regard parents as better informants because they see their children across so many diverse settings. If teachers should detect children's psychosocial problems better than parents due to experience, then experienced parents with several children (older siblings of the target child) should be more similar to teachers in their ratings than inexperienced parents, even if parents rarely care for as many children as preschool teachers routinely do. To test this, we gathered information about the age and the number of children parents had. The number of older siblings was entered in the GLM model

(Table 4) but did not have any effect on the model. Neither teachers' nor parents' experience with children enhanced agreement.

### Child Problem Severity Related to Parent–Teacher Disagreement

The more severe the child's problems were, the poorer the agreement between parent and teacher tended to be (see Table 6). When parents rated children as having a high problem severity, 38.8% of teachers rated them as having moderate problems. Just as many teachers, though (39%), rated the children as having no or few problems. It is not difficult to picture both referral problems and diagnostic uncertainty when parents and teachers disagree on three out of four children who have a high problem severity.

### Parental Characteristics that Contributed to Parent–Teacher Disagreement

Increased disagreement between parents and teachers was related to elevated parental stress (Table 4). Stress as a determinant is in line with earlier studies where it has been related to mothers' over-reporting of behavioral problems [68, 69]. The interpretation has been that stressed mothers tend to perceive their children's behavior as more negative [70]. Our study confirms this, since increased self-reported stress was strongly associated with more parent–teacher disagreement. The parents, compared to the teachers, then reported more child problems than usual.

Parental depression, however, was not a predictor of disagreement in our sample. The mean score for depressive symptoms in our sample was 5.6, which indicates a low proportion of parents actually being depressed. Clearly more work with community samples of preschoolers, preferably in more countries, needs to be conducted before one can conclude whether the propensity to distort because of depression is mainly a characteristic of clinical samples. None of the other parental characteristics such as deviant personality traits or anxiety predicted CBCL-TRF discrepancy. This again may possibly be due to the low-risk population sample.

### Cross-Cultural Issues

Since Danish teachers did not rate preschoolers' internalizing problems any lower than parents did [63], the findings in our study of teachers' low scores on internalizing problems may be a more culture-specific phenomenon. Yet, a Dutch study with a preschool sample found, as in our study, that the largest difference between parent–teacher informers was with internalizing problems [71]. Perhaps Danish teachers have better training in children's mental health and therefore are more sensitive to internalizing behavior. A cross-national study that considers teachers' training and perceptions of children might prove to be interesting.

However, the finding that the teacher–child conflict increased discrepancies on all three CBCL/TRF scales is more likely to be relevant beyond a Scandinavian population. The Norwegian preschool system is rather lax, permissive, and play-focused. In other countries, stricter and more academically challenging preschools with greater demands on children and teachers might possibly generate more perception biasing teacher–child conflict. In line with this, conflict in the classroom has been associated with US teachers' authoritarian attitudes [72].



In the current study, the percentage of childcare personnel with a relevant education was high among those that had completed questionnaires (83.7%) and considerably higher than that in most Norwegian childcare centers and conceivably in other countries. The caregivers or teachers had known the child an average of 18.8 months, which is longer compared to other countries. However, neither education nor prior experience, nor knowledge of the child affected the discrepancy scores in this study. Nevertheless, in a previous study of a large US sample of school-age children [15], teachers' familiarity with a child enhanced the parent–teacher agreement. These diverging results may be due to more variance in how familiar preschool teachers and caregivers are with the children in the US compared to in Norway. In Norwegian daycare centers, one strives to obtain as much continuity in the teacher–child relationship as possible, trying to avoid changing a child's preschool teacher or assigned caretaker. In addition, we requested that the preschool teacher or caretaker who knew the child the best be the one who completed the TRF. Both of these circumstances may explain why the variable of familiarity with the child did not affect discrepancy in the current study.

### Limitations

Despite various strengths of the current study, including a large, representative sample of a known community, a possible selection bias cannot be ruled out since 10% of the preschool teachers did not return the forms and an additional 11% had incomplete data.

Longitudinal data would enable a better evaluation of the information obtained from parents and teachers by investigating their differential predictive power vis-a-vis later problem behavior, as the studies of Ferdinand et al. [9, 10] and Verhulst et al. [8] have demonstrated.

In the present study, the factors that were most related to parent–teacher discrepancies were conflict in the teacher–child relationship and, to a lesser degree, perceived stress in the parent role. These measures most likely also reflect child psychopathology as manifested in different relationships and contexts. Thus, what the STRS (Student–Teacher Relationship Scale) and PSI (Parenting Stress Index) measure may have just as much to do with the characteristics of the child as the informant. The hitherto-undocumented hypothesis that context-specific child behavior accounts for most of the disagreement between parents and teachers could not be directly addressed in this study.

### Summary

The necessity of better understanding of informant discrepancies in child mental health has been called for by several researchers. In this study, teacher factors were included as possible determinants of parent–teacher (dis)agreement on preschoolers' psychosocial problems. A total of 732 4-year olds from a Norwegian community sample were assessed with ASEBA. Furthermore, teachers reported on their education, experience and relationship to the child (STRS). Parental stress and psychopathology were also measured. Overall, teachers rated children considerably lower than parents, especially on internalizing problems. When teachers did rate more child problems, this was strongly linked to teachers' perceptions of a teacher–child relationship characterized by conflict. This factor contributed more than the others to disagreement. The highest agreement occurred on boys' externalizing problems. Compared to parents, however, girls' behavior was rated much lower by teachers than boys' behavior. Neither teachers' nor parents' increased

experience with children enhanced agreement. Parents and teachers disagreed more on children whom they rated to be troubled than on non-problematic children. Parental characteristics that contributed to prediction of disagreement were parenting stress. The main findings of this study on parent–teacher disagreement showed teacher determinants such as teacher–child conflict, under-reporting of internalizing problems and a possible same-gender child preference and discussed them as possible reporting biases.

**Acknowledgments** We wish to thank the families and teachers that generously volunteered their time to participate. This research was funded in part by grants 170449/V50, 190622/V50, and 175309/V50 from the Research Council of Norway and grant 4396 from the Liaison Committee between the Central Norway Regional Health Authorities and the Norwegian University of Science and Technology.

**Conflict of interest** None.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

## References

1. De Los Reyes A, Kazdin AE (2004) Measuring informant discrepancies in clinical child research. *Psychol Assess* 16:330–334
2. Carter AS, Briggs-Gowan MJ, Davis NO (2004) Assessment of young children's social-emotional development and psychopathology: recent advances and recommendations for practice. *J Child Psychol Psychiatry* 45:109–134
3. De Los Reyes A, Goodman KL, Kliewer W, Reid-Quinones K (2008) Whose depression relates to discrepancies? Testing relations between informant characteristics and informant discrepancies from both informants' perspectives. *Psychol Assess* 20:139–149
4. Achenbach TM, McConaughy SH, Howell CT (1987) Child/adolescent behavioral and emotional problems: implications of cross-informant correlations for situational specificity. *Psychol Bull* 101: 213–232
5. Kolko DJ, Kazdin AE (1993) Emotional/behavioral problems in clinic and nonclinic children: correspondence among child, parent and teacher reports. *J Child Psychol Psychiatry* 34:991–1006
6. Wolraich ML, Lambert EW, Bickman L (2004) Assessing the impact of parent and teacher agreement on diagnosing attention-deficit hyperactivity disorder. *J Dev Behav Pediatr* 25:41–47
7. Sourander A, Piha J (1997) Parent, teacher and clinical ratings on admission to child psychiatric, inpatient treatment: a study of cross-informant correlations using the CBCL, Rutter scales and the CGAS. *Nord J Psychiatry* 51:365–370
8. Verhulst FC, Dekker MC, van der Ende J (1997) Parent, teacher and self-reports as predictors of signs of disturbance in adolescents: whose information carries the most weight? *Acta Psychiatr Scand* 96:75–81
9. Ferdinand RF, van der Ende J, Verhulst FC (2007) Parent-teacher disagreement regarding psychopathology in children: a risk factor for adverse outcome? *Acta Psychiatr Scand* 115:48–55
10. Ferdinand RF, Hoogerheide KN, van der Ende J, Visser JH, Koot HM, Kasius MC, Verhulst FC (2003) The role of the clinician: three-year predictive value of parents' teachers' and clinicians' judgment of childhood psychopathology. *J Psychol Psychiatry* 44:867–876
11. Verhulst FC, Koot HM, van der Ende J (1994) Differential predictive value of parents' and teachers' reports of children's problem behaviors: a longitudinal study. *J Abnorm Child Psychol* 22:531–546
12. McConaughy SH, Stanger C, Achenbach TM (1992) Three years course of behavioral/emotional problems in a national sample of 4- to 16-year-olds: agreement among informants. *J Am Acad Child Adolesc Psychiatry* 31:932–940
13. Kraemer HC, Measelle JR, Ablow JC, Essex MJ, Boyce WT, Kupfer DJ (2003) A new approach to integrating data from multiple informants in psychiatric assessment and research: mixing and matching contexts and perspectives. *Am J Psychiatry* 160:1566–1577
14. Konold TR, Pianta RC (2007) The Influence of informants on ratings of children's behavioral functioning. A latent variable approach. *J Psychoeduc Assess* 25:222–236

15. Zahner G, Daskalakis C (1998) Modeling sources of informant variance in parent and teacher ratings of child psychopathology. *Int J Method Psych* 7:3–16
16. Deng S, Liu X, Roosa MW (2004) Agreement between parent and teacher reports on behavioral problems among Chinese children. *J Dev Behav Pediatr* 25:407–414
17. Kerr DCR, Lunkenheimer ES, Olson SL (2007) Assessment of child problem behaviors by multiple informants: a longitudinal study from preschool to school entry. *J Child Psychol Psychiatry* 48:967–975
18. Verhulst FC, Akkerhuis GW (1989) Agreement between parent's and teacher's ratings of behavior/emotional problems of children aged 4–12. *J Child Psychol Psychiatry Allied Disciplines* 30:123–136
19. Richter JE (1992) Depressed mothers as informants about their children: a critical review of the evidence for distortion. *Psychol Bull* 112:485–499
20. Berg-Nielsen TS, Vikan A, Dahl AA (2003) When adolescents disagree with their mothers: CBCL-YSR discrepancies related to maternal depression and adolescent self-esteem. *Child Care Health Dev* 29:207–213
21. Chi TC, Hinshaw SP (2002) Mother-child relationships of children with ADHD: the role of maternal depressive symptoms and depression-related distortions. *J Abnorm Child Psychol* 30:387–400
22. De Los Reyes A, Kazdin AE (2005) Informant discrepancies in the assessment of childhood psychopathology. A critical review, theoretical framework and recommendations for further study. *Psychol Bull* 131:483–509
23. Foley D, Rutter M, Pickles A, Angold A, Maes H, Silberg J, Eaves L (2004) Informant disagreement for separation anxiety disorder. *J Am Acad Child Adolesc Psychiatry* 43:452–460
24. Goodman R (1997) The strengths and difficulties questionnaire: a research note. *J Child Psychol Psychiatry* 38:581–586
25. Goodman R, Ford T, Simmons H, Gatward R, Meltzer H (2000) Using the Strengths and Difficulties Questionnaire (SDQ) to screen for child psychiatric disorders in a community sample. *Br J Psychiatry* 177:534–539
26. Statistics Norway (2010) Figures on Trondheim municipality. Retrieved 05.04.2011 from <http://www.ssb.no/english/municipalities/1601>
27. Pianta RC (1996) Manual and scoring guide for the Student-Teacher Relationship Scale. University of Virginia, Charlottesville
28. Pianta RC (2001) STRS student-teacher relationship scale. Professional manual: PAR: Psychological Assessment Resources, Inc
29. Achenbach TM, Rescorla LA (2000) Manual for the ASEBA Preschool Forms and Profiles: an integrated system of multi-informant assessment. University of Vermont Department of Psychiatry, Burlington
30. Achenbach TM, Becker A, Dopfner M, Heiervang E, Roessner V, Steinhausen HC, Rothenberger A (2008) Multicultural assessment of child and adolescent psychopathology with ASEBA and SDQ instruments: research findings, applications, and future directions. *J Child Psychol Psychiatry* 49:251–275
31. Goodman R (2001) Psychometric properties of the strengths and difficulties questionnaire. *J Am Acad Child Adolesc Psychiatry* 40:1337–1345
32. Crone MR, Vogels AGC, Hoekstra F, Treffers PDA, Reijneveld SA (2008) A comparison of four scoring methods based on the parent-rated Strengths and Difficulties Questionnaire as used in the Dutch preventive child health care system. *BMC Public Health* 8:106
33. Obel C, Heiervang E, Rodriguez A, Heyerdahl S, Smedje H, Sourander A et al (2004) The strengths and difficulties questionnaire in the Nordic countries. *Eur Child Adolesc Psychiatry* 13:32–39
34. Heiervang E, Stormark KM, Lundervold AJ, Heimann M, Goodman R, Posserud MB et al (2007) Psychiatric disorders in Norwegian 8 to 10-year-olds: an epidemiologic survey of prevalence, risk factors, and service use. *J Am Acad Child Adolesc Psychiatry* 46:438–447
35. Sanne B, Torsheim T, Heiervang E, Stormark KM (2009) The strengths and difficulties questionnaire in the Bergen child study: a conceptually and methodically motivated structural analysis. *Psychol Assess* 23:352–364
36. Bérubé RL, Achenbach TM (2005) Bibliography of published studies using ASEBA instruments: 2005 edition. Research Center for Children, Youth, and Families, University of Vermont, Burlington VT
37. Rescorla LA (2005) Assessment of young children using the Achenbach System of Empirically Based Assessment (ASEBA). *Ment Retard Dev Disabil Res Rev* 11:226–237
38. Bingham CR, Loukas A, Fitzgerald HE, Zucker RA (2003) Parental ratings of son's behavior problems in high-risk families: convergent validity, internal structure, and interparent agreement. *J Pers Assess* 80:237–251
39. Xing Tan T, Dedrick RF, Marfo K (2007) Factor structure and clinical implications of child behavior checklist/1.5–5 ratings in a sample of girls adopted from China. *J Pediatric Psychol* 32:807–818

40. Ivanova MY, Achenbach TM, Rescorla LA, Harder VS, Ang RP, Bilenberg N et al (2010) Preschool psychopathology reported by parents in 23 societies: testing the seven-syndrome model of the Child Behavior Checklist for ages 1.5–5. *J Am Acad Child Adolesc Psychiatry* 49:1215–1224
41. Hutzelmeyer-Nickels A, Noterdaeme M (2007) Usefulness of the Child Behavior Checklist (CBCL 1 1/2–5) in the assessment of preschool children with developmental problems. *Prax Kinderpsychol Kinderpsychiatr* 56:573–588
42. Denner S, Schmeck K (2005) Emotional and behavioral disorders at preschool age—results of a study in kindergarten children in Dortmund using the Caregiver-teacher report From C-TRF 1 1/2 -5. *Zeitschrift für Kinder und Jugendpsychiatrie und Psychotherapie* 33:307–317
43. Cheng LJ, Leung PW (2011) The application of the preschool Child Behavior Checklist and the Caregiver-teacher Report Form to Mainland Chinese children: syndrome structure, gender differences, country effects, and inter-informant agreement. *J Abnorm Child Psychol* 39:251–264
44. Drugli M, Larsson B, Clifford G, Fossum S (2007) Pervasive and non-pervasive conduct problems in a clinic sample aged 4–8 years: child, family and day-care/school factors. *Scand J Educ Res* 51:547–559
45. Drugli M, Larsson B (2006) Children aged 4–8 years treated with parent training and child therapy because of conduct problems: generalisation effects to day-care and school settings. *Eur Child Adolesc Psychiatry* 15:392–399
46. Pianta RC (1994) Patterns of relationships between children and kindergarten teachers. *J Sch Psychol* 32:15–31
47. Solheim E, Wichstrom L, Berg-Nielsen, TS (2011) The Student-Teacher-Relationship Scale: a valid instrument in preschool samples? *J Psychoeduc Assess* (in press)
48. Doumen S, Verschueren K, Buyse E, Germeijs V, Luyckx K, Soenens B (2008) Reciprocal relations between teacher-child conflict and aggressive behavior in kindergarten: a three-wave longitudinal study. *J Clin Child Adolesc Psychol* 37:588–599
49. Palermo F, Hanish L, Martin C, Fabes R, Reiser M (2007) Preschoolers' academic readiness: what role does the teacher-child relationship play? *Early Child Res Q* 22:407–422
50. Rudasill KM, Rimm-Kaufman SE, Justice LM, Pence K (2006) Temperament and language skills as predictors of teacher-child relationship quality in preschool. *Early Educ Dev* 17:271–291
51. Abidin RR (1990) Parenting stress index clinical manual. Pediatric Psychology Press, Charlottesville
52. Ostberg M, Hagekull B, Wettergreen S (1997) A measure of parental stress in mothers with small children: dimensionality, stability and validity. *Scand J Psychol* 38:199–208
53. Beck AT, Steer RA, Ball R, Ranieri WF (1996) Comparison of Beck Depression Inventories-IA and -II in psychiatric outpatients. *J Pers Assess* 67:588–597
54. Richter P, Werner J, Heerlein A, Kraus A, Sauer H (1998) On the validity of the Beck Depression Inventory. *Psychopathology* 31:160–168
55. Arnarson PO, Olason DP, Smari J, Sigurdsson JF (2008) The Beck Depression Inventory Second Edition (BDI-II): psychometric properties in Icelandic student and patient populations. *Nord J Psychiatry* 62:360–365
56. Beck AT, Brown G, Epstein N, Steer RA (1988) An inventory for measuring clinical anxiety—psychometric properties. *J Consult Clin Psychol* 56:893–897
57. Ottosson H, Bodlund O, Ekselius L, von Knorring L, Kullgren G, Lindstrom E et al (1995) The DSM-IV and ICD-10 Personality Questionnaire (DIP-Q)-construction and preliminary validation. *Nord J Psychiatry* 49:285–291
58. Gadermann A, Guhn M, Zumbo BD (2008) An empirical comparison of Cronbach's alpha with ordinal reliability coefficients alpha and theta. *Int J Psychol* 43:55
59. Zumbo BD, Gadermann AM, Zeisser C (2007) Ordinal versions of coefficients alpha and theta for Likert rating scales. *J Mod Appl Stat* 6:21–29
60. Ottosson H, Bodlund O, Ekselius L, Grann M, von Knorring L, Kullgren G et al (1998) DSM-IV and ICD-10 personality disorders: a comparison of a self-report questionnaire (DIP-Q) with a structured interview. *Eur Psychiatr* 13:246–253
61. Ottosson H, Grann M, Kullgren G (2000) Test-retest reliability of a self-report questionnaire for DSM-IV and ICD-10 personality disorders. *Eur J Psychol Assess* 16:53–58
62. Hamre BK, Pianta RC, Downer JT, Mashburn AJ (2008) Teachers' perception of conflict with young students: looking beyond problem behaviors. *Social Dev* 17:115–136
63. Kristensen S, Henriksen TB, Bilenberg N (2010) The Child Behavior Checklist for ages 1.5–5 (CBCL/1 1/2-5): assessment and analysis of parent- and caregiver-reported problems in a population-based sample of Danish preschool children. *Nord J Psychiatry* 64:203–209
64. Stanger C, Lewis M (1993) Agreement among parents, teachers and children on internalizing and externalizing behavior problems. *J Clin Child Psychol* 22:107–115

65. Egger HL, Angold A (2006) Common emotional and behavioral disorders in preschool children: presentation, nosology, and epidemiology. *J Child Psychol Psychiatry* 47:313–337
66. Wichstrom L, Angold A, Berg-Nielsen TS, Egger HL, Solheim E, Sveen TH (2012) Prevalence and correlates of psychiatric disorders in preschoolers. *J Child Psychol Psychiatry* (in press)
67. Mesman J, Koot H (2000) Child-reported depression and anxiety in preadolescence: I. Associations with parent- and teacher-reported problems. *J Am Acad Child Adolesc Psychiatry* 39:1371–1378
68. Bonde E, Obel C, Nedergard NJ, Thomsen PH (2004) Social risk factors as predictors for parental report of deviant behaviour in 3-year old children. *Nord J Psychiatry* 58:17–23
69. Seiffge-Krenke I, Kollmar F (1998) Discrepancies between mothers' and fathers' perception of sons' and daughters' problem behaviour: a longitudinal analysis of parent-adolescent agreement on internalising and externalising problem behaviour. *J Child Psychol Psychiatry* 39:687–697
70. Crnic KA, Gaze C, Hoffman C (2005) Cumulative parenting stress across the preschool period: relations to maternal parenting and child behavior at age 5. *Inf Child Dev* 14:117–132
71. Grietens H, Onghena P, Prinzie P, Gadeyne E, Van Assche V, Ghesgüière P, Hellinckx W (2004) Comparison of mothers', fathers' and teachers' reports on problem behavior in 5- to 6-year-old children. *J Psychopathol Behav* 26:137–146
72. Pianta RC, Howes C, Burchinal M, Bryant D, Clifford R, Early C et al (2005) Features of pre-kindergarten programs, classrooms and teachers: do they predict observed classroom quality and child-teacher interactions? *Appl Dev Sci* 9:144–159