Received: 26 March 2015 | Revised: 2 August 2017 | Accepted: 8 August 2017

DOI: 10.1002/mpr.1589

WILEY

ORIGINAL ARTICLE

Understanding parent–teacher agreement of the Strengths and Difficulties Questionnaire (SDQ): Comparison across seven European countries

Sabrina Cheng[1,2] | Katherine M. Keyes[1] | Adina Bitfoi[3] | Mauro Giovanni Carta[4] | Ceren Koç[5] | Dietmar Goelitz[6] | Roy Otten[7] | Sigita Lesinskiene[8] | Zlatka Mihova[9] | Ondine Pez[10] | Viviane Kovess-Masfety[10]

[1] Department of Epidemiology, Mailman School of Public Health, Columbia University, New York, NY, USA

[2] New York University School of Medicine, New York, NY, USA

[3] The Romanian League for Mental Health, Bucharest, Romania

[4] Department of Public Health, Center of Liaison Psychiatry, Cagliari University, Cagliari, Italy

[5] Yeniden Health and Education Society, Istanbul, Turkey

[6] Friedrich-Alexander University Erlangen-Nurenberg, Erlangen, Germany

[7] Faculty of Social Sciences, Radboud University Nijmegen, Nijmegen, The Netherlands

[8] Clinic of Psychiatry, Faculty of Medicine, University of Vilnius, Vilnius, Lithuania

[9] Department of Psychology, New Bulgaria University, Sophia, Bulgaria

[10] Department MéTis, École des Hautes Études en Santé Publique, Paris Descartes University, Paris, France

Correspondence
Viviane Kovess-Masfety, Department MéTis
20 avenue George Sand, La Plaine Saint Denis 93210, France
Email: viviane.kovess@ehesp.fr

Funding information
European Union, Grant/Award Number: 2006336

Abstract

Assessments of child psychopathology are often derived from parental and teacher reports, yet there is substantial disagreement. This study utilized data from 7 European countries to examine parent–teacher agreement and possible explanatory factors for parent–teacher disagreement such as child and family characteristics, parenting dimensions, and maternal distress were explored. Parent–teacher agreement of the Strengths and Difficulties Questionnaire were assessed using a cross-sectional survey of 4,894 school aged children 6–11 from the School Children Mental Health Europe Project. Parent–teacher agreement was low to moderate (Pearson correlation ranging from .24 (Prosocial) to .48 (Hyperactivity) for the 5 subscales across 7 countries); kappa coefficient ranged from .01 (Turkey) to .44 (Italy) for internalizing problems and .19(Romania) to .44(Italy) for externalizing problems. Child's gender and age, mother's employment status, single parent home, number of children in household, and selected parenting dimension were found to be explanatory of informant disagreement. This study not only serves to advance our understanding of parent–teacher agreement of the Strengths and Difficulties Questionnaire in 7 European countries but provides a novel approach to examining the factors that contribute to informant disagreement.

KEYWORDS

externalizing behavior, informant discrepancies, internalizing behavior, SDQ, young children

1 | INTRODUCTION

Assessment of child psychopathology is critical for early identification and treatment. Among young children, multiple informants are almost universally used for assessment and diagnosis, including parents and teachers reports (De Los Reyes, Thomas, Goodman, & Kundey, 2013). Although parents and teachers reports are critical to understanding child mental health, studies have consistently shown that there is low to modest agreement across informants (Achenbach, McConaughy, & Howell, 1987; Duhig, Renk, Epstein, & Phares, 2000). The Achenbach et al. (1987) meta-analysis consisting of 119 studies found that different informants' (e.g., parent, child, and teacher) ratings of social, emotional, or behavioral problems in children are often discrepant, with mean Pearson correlation coefficient of .22–.28 (Achenbach et al., 1987). Even among parents (mother–father), there were only moderate agreements (correlation of .60; Achenbach

et al., 1987). Subsequent studies have also found similar results (De Los Reyes et al., 2015; De Los Reyes & Kazdin, 2005; Rescorla et al., 2014; Stone, Otten, Engels, Vermulst, & Janssens, 2010). The De Los Reyes et al. (2015) meta-analysis consisting of 341 studies also found that for internalizing and externalizing behaviors, the cross-informant correspondence was .25 and .30, respectively.

In the last decade, there has been a shift from establishing and describing informant discrepancies to determining the values and benefits of the discrepancies itself (De Los Reyes, 2011; De Los Reyes & Kazdin, 2004; De Los Reyes & Kazdin, 2005). Since there is no cost-effective biological or behavioral marker that can be used to diagnose mental health disorders in children, understanding the constructs and factors that influence multi-informant discrepancies might improve evidence-based assessment of psychopathology in children and provide more detailed guidelines on how to monitor and evaluate treatment planning. Generally, literature has shown that informant pair, problem behavior type (internalizing compared with externalizing problems), and measurement method influence the magnitude of informant discrepancies (Achenbach et al. 1987; De Los Reyes & Kazdin, 2005; De Los Reyes et al., 2015; Duhig et al., 2000; Guion, Mrug, & Windle, 2009). Child gender and age have shown mixed findings including two meta-analyses showing no gender effect (Achenbach et al. 1987; De Los Reyes & Kazdin, 2005; Duhig et al., 2000; Harvey, Fischer, Weieneth, Hurwitz, & Sayer, 2013). While Achenbach et al. (1987) found child's age (age 12- to 16-year-olds compared with 6- to 11-year-olds) to be associated with informant discrepancies, De Los Reyes et al. (2015) did not find child age to be related to informant discrepancies. Literature has also found that socioeconomic status and family structure such as family status (i.e., divorced vs. intact families), number of children in the family, mother's age, and education and employment status can influence the magnitude of informant discrepancies (Harvey et al., 2013; Jensen, Xenakis, Davis, & Degroot, 1988; Michels et al., 2013; Stone, Speltz, Collett, & Werler, 2013; Van Roy, Groholt, Heyerdahl, & Clench-Aas, 2010).

Informant discrepancy can also be related to parenting and parent psychopathology. Specifically, maternal depression, maternal stress, and global measures of psychopathology are associated with skewed perceptions of their children's problems that are often not confirmed by children or teachers reports (Briggs-Gowan, Carter, & Schwab-Stone, 1996; De Los Reyes & Kazdin, 2006; Ehrlich, Cassidy, Lejuez, & Daughters, 2013; Richters, 1992). In terms of parenting dimensions, literature has found that positive parent–child relationship and less parenting stress are associated with less informant discrepancies (Chi & Hinshaw, 2002; Fung & Lau, 2010; Pelton, Steele, Chance, & Forehand, 2001; Treutler & Epkins, 2003; Van der Oord, Prins, Oosterlaan, & Emmelkamp, 2006). Authoritative parenting (characterized by responsive and nurturing parenting styles) has been found to be associated with lower ratings of emotion problems (i.e., anger, sadness, and anxiety) by parents than by children self-reports (Michels et al., 2013).

Informants from different backgrounds might vary in what they perceive to be problematic behaviors that warrant concerns. Understanding the differences or similarities in informant disagreement across countries can improve decisions on how data from multiple informants can be used. Studies across different countries have consistently shown that agreement between multiple informants is low but the amount of discrepancies varies across countries (Guion et al., 2009; Harvey et al., 2013; Rescorla et al., 2013; Rescorla et al., 2014; Van der Ende & Verhulst, 2005). A recent review on parent–teacher agreement across 21 countries using the Child Behavior Checklist (CBCL) and Teacher Report Form (TRF) found that although there are numerous similarities, the magnitude of parent–teacher agreement varied across countries (Rescorla et al., 2014). In the study by Rescorla and colleagues, they found kappa's ranging from .03 (Thailand and Singapore) to .35 (Denmark) for internalizing behavior problems and from .03 (Thailand) to .35 (Jamaica) for externalizing behavior problems. These country variations might impact mental health diagnoses and cross-cultural research on disorder prevalence.

The majority of studies examining parent–teacher discrepancies have focused on the CBCL and TRF (De Los Reyes et al., 2015; Rescorla et al., 2014). While the CBCL is often used for in-depth assessment, the SDQ offers an alternative that is more suitable for screening purposes; the brevity and ease of the measure have seen increase usage worldwide, especially for research purposes (Stone et al., 2010). Stone et al. (2010) conducted a systematic review evaluating results from 48 studies on the reliability and validity of the parent and teacher SDQ for 4- to 12-year-olds in over 20 countries, concluding that the psychometric properties of the SDQ are strong and should continue to be used as a screening instrument. While Stone et al. (2010) has examined parent–teacher agreement of the SDQ in their review, they have also stressed the importance of further understanding multi-informant and the role that culture plays in the distribution and expression of psychosocial problems in society, thus making it imperative to study the SDQ in different countries.

This study examined parent–teacher agreement of the SDQ for 4,894 school-age children in seven diverse European countries. Possible explanatory factors for informant disagreement such as child, maternal, and family characteristics, parenting dimensions, and maternal distress were explored. The purpose of this study is to answer the following questions:

1. Do parent–teacher agreements using the SDQ vary between different problem behaviors and do they vary between countries? Based on previous research, we hypothesized that parent–teacher agreement would be higher for externalizing behavior than internalizing behavior and that agreement will vary across countries.
2. Do parent–teacher agreements vary by child, maternal, and family characteristics, parenting dimensions, maternal distress, and European region? Based on previous research, we hypothesized that parent–teacher agreement will be varied by child age, maternal education status, age and employment status, family structure, parenting dimension, and maternal stress.

This study will contribute to the current literature in several ways. First, this study examines parent and teacher agreement in the

assessment of child psychopathology using the SDQ in seven European countries including Bulgaria, Germany, Italy, Lithuania, Netherlands, Romania, and Turkey. Second, this will contribute to a growing area of literature examining the impact of child, maternal and family characteristics, parenting dimension, and parental psychopathology on informant discrepancies. Furthermore, understanding factors that would provide insight to why variability across parent–teacher reports occurs can potentially guide researchers and clinicians when they are faced with discrepant data (Harvey et al., 2013).

2 | METHODS

2.1 | Samples and procedure

The School Child Mental Health Europe project was a cross-sectional survey of school children aged 6 to 11 years and their parents and teachers in seven European countries (Bulgaria, Germany, Italy, Lithuania, Netherlands, Romania, and Turkey) in 2010. A two-level random selection procedure was utilized. In each grade, classes were randomly selected from randomly selected schools in each country, and six children were randomly selected from each class. Approximately 48 children were randomly selected from each school with the exception of Netherlands, where more children were selected due to lower participation of schools. Enrollment was based on passive consent; parents received an informational letter and consent form. If the parents did not mail the school a consent form stating their refusal to participate, the child was included. Of invited participants, 72.2% participated in the survey from Western Europe and 61.3% participated in the survey from Eastern Europe. Details of methodology are reported elsewhere (Pez et al. 2011).

Only subjects that had information reported from both parent and teacher were included in this study. In most cases, the corresponding parental respondent was the mother (86.7%). In order to reduce heterogeneity, we restricted these analyses to mother's report only. The total sample size for this study was 4,894 subjects. For the 4,894 subjects in this study, 911 subjects were from Bulgaria (18.61%), 370 subjects were from Germany (7.56%), 687 subjects were from Italy (14.04%), 995 subjects were from Lithuania (20.33%), 531 subjects were from Netherlands (10.85%), 932 subjects were from Romania (19.04%), and 468 subjects were from Turkey (9.56%).

2.2 | Instruments

2.2.1 | Child mental health

Parents and teachers reported on child mental health using the SDQ (Goodman, 1997). SDQ is a widely used and reliable self-administered psychopathology screening for children aged 4–16 years old. Parents and teachers reported on their children's behaviors over the past 6 months, with a 3-point scale ranging from 0 = not true to 2 = certainly true. It contains 25 questions generating five subscales: emotional problems, hyperactivity and inattention, conduct problems, peer relationship difficulties, and prosocial behaviors. Broader internalizing and externalizing subscales were also created. Externalizing problem was defined as either greater than the cutoff on conduct problems (greater or equal to 4) and/or attention deficit hyperactivity disorder symptoms (greater or equal to 7). Internalizing problem was defined as being greater than the cutoff on emotional problems (greater or equal to 5 for parent and greater or equal to 6 for teacher). These cutoffs were based on the abnormal cutoff of 90 percentile from Goodman (1997) epidemiological study. When using these cutoffs, we found that 84.3%–93% of children in this sample are within the normal range (see Kovess et al., 2015 for details and prevalence estimates).

As part of School Child Mental Health Europe substudy, instruments were validated against clinician's judgment in the participating countries and were tested and retested for reliability and consistency. In this study, the conduct problem behavior, peer problem, and prosocial problem showed relatively low reliability for parents (α = .59, .49, and .64, respectively) consistent with those found in Stone et al. (2010; α = .58, .53, and .67, respectively). Reliability for emotional problems and hyperactivity scales showed adequate reliability for parents (α = .67, and .72, respectively). The reliability for teacher ranged from α = .75 to .84 for emotional problems, conduct problems, hyperactivity, and prosocial behavior scale; and for peer relationship α = .58. Despite low internal consistency estimates, all subscales of the SDQ were included in this study to provide an opportunity to compare our results with other relevant studies.

An impact supplement was also included. Impact score was calculated if the child had a perceived difficulty in any of the three problem areas, namely emotions, concentration, or behavior. For parents ratings, impact score was calculated with 4 items (using a "0012" scale: 0 = not at all/only a little, 1 = quite a lot, 2 = a great deal) measuring social interference and distress, for a maximum of 10 points. For teachers, there were only 2 social interference items, for a maximum impact score of 6. A total impact score greater than 1 is defined as abnormal, as is consistent with prior literature (see Goodman, 1999 for details). Any child with an impact score greater than 1 was considered to "have impact." Thus, for example, a child with a "perceived difficulty" in concentration (i.e., a score above the cutpoint, a rating of 2) who was rated 2 = a great deal on any of the 4 (2 for teachers) social interference items or the distress item would be considered to have impact.

Impact score was used to obtain additional information about the severity of the problem behavior. The impact score conceptually correspond to a score of 60 or less on the Children's Global Assessment Scale, which is a level that has been shown to be a good indicator of definite caseness (Goodman, 1999). Externalizing and internalizing problem caseness was determined as having the respective behavior problem and scoring above 1 in the impact scale. Parent–teacher agreement/disagreement on internalizing and externalizing caseness was also defined. If parent and teacher agreed on externalizing and internalizing caseness, a score of 0 was given; and if parent and teacher disagreed on externalizing and internalizing caseness, a score of 1 was given.

2.2.2 | Child, maternal, and family characteristics

We included the following child, maternal, and family characteristics in the analyses as possible explanatory variables for assessing

interinformant disagreement: gender, child age, mother age, maternal education, employment status, number of children in household, and single motherhood.

2.2.3 | Parenting dimensions

The Parenting Scale (TPS) and Parent Behaviors and Attitudes Questionnaire were used to measure five domains of parenting (caring, autonomous, laxness, overreactivity, and verbosity).

TPS (Arnold, O'Leary, Wolff, & Acker, 1993) is a 30-item measure of dysfunctional discipline practices that measures the probability that a parent uses particular discipline strategies. Three discipline styles were measured: Laxness (permissive discipline), Overreactivity (displays of anger, meanness, and irritability), and Verbosity (lengthy verbal responses or reliance on talking). TPS uses a Likert scale from 1 to 7. In this study, we excluded the verbosity scale (α = .50) because of low reliability. Laxness and Overreactivity subscales had adequate internal consistencies in the current study (α = .70).

The Parent Behaviors and Attitudes Questionnaire (Bergeron, Valla, & Breton, 1992) includes 23 items assessed on a 4-point Likert scale (all of the time, most of the time, sometimes, and never). For the purpose of this study only Autonomy and Caring scales were used and were based on seven and eight items, respectively. Items were summed, and cutoff points were created based on being more than one standard deviation above the mean. Both Autonomy and Caring subscales showed adequate internal consistencies in this study (α = .77 and .72).

2.2.4 | Parent psychopathology

Maternal psychopathology information was also obtained as a possible explanatory variable for interinformant disagreement. Maternal psychopathology was assessed using the MH-5 scale of the Short-Form-36 (Aaronson et al., 1992) with the following scales: Psychological distress (MH-5), Vitality (VT4), and Role emotional (RE). Each scale is scored from 0 to 100, the highest functional status level at 100 and below 57 as distress (Wadsworth, Corley, Hewitt, & DeFries, 2001).

2.3 | Statistical analyses

2.3.1 | Parent–teacher informant agreement using the SDQ

First, we wanted to establish the patterns of parent and teacher report of child behavior by countries. Basic descriptive were performed to obtain means and standard deviations. We evaluated whether the SDQ scores varied by country using PROC GLIMMIX in SAS. We fitted the model with the SDQ score as dependent and random country effects.

Next, we wanted to examine the degree of agreement by parent and teacher. Parent and teacher agreement on the SDQ was measured continuously and categorically, using Pearson correlation coefficient and kappa statistic. Pearson correlation (r) was used to examine the magnitude of association among parent and teacher reports on total difficulties and five subscales scores of the SDQ. For kappa statistic, a two by two comparison of parent and teacher reported internalizing and externalizing caseness were examined. Kappa coefficient less than 0 indicated a less than chance agreement, .01–.20 indicated a slight agreement, .21–.41 indicated a fair agreement, and .41–.60 indicated a moderate agreement, .61–.80 indicated a substantial agreement, and .81–.99 indicated an almost perfect agreement (Viera & Garrett, 2005).

We also wanted to test whether parent–teacher agreement on externalizing and internalizing caseness varied by countries. Using the dichotomous variable of parent–teacher agreement on externalizing and internalizing caseness, chi-square was conducted to test for country (group) differences. For chi-square that showed overall country differences, post hoc group comparisons were further carried out to understand specific country differences.

2.3.2 | Factors that may influence agreement

Logistic regressions were then used to examine the association of child and family characteristics, country region, parenting styles, and maternal distress on interinformant disagreement on externalizing and internalizing caseness. Models were estimated by externalizing and internalizing caseness disagreement and expressed as odds ratios (ORs) with 95% confidence interval. First, externalizing and internalizing caseness disagreement was regressed on the predictors individually to determine unadjusted ORs. Significant predictors ($p \leq .001$ from the unadjusted analyses) were then simultaneously entered in the final logistic regression analyses. This was done in order to understand which predictors contributed above and beyond other significant predictors and to control for the relationship among the predictors. A p value of .001 was chosen due to the expected inflation of the type I error from multiple statistical testing.

3 | RESULTS

3.1 | Descriptive

Demographic data on study sample is presented in Table 1. Child age ranged from 5 to 13 years with a mean of $M = 8.67$ years ($SD = 1.36$) and maternal age ranged from 24 to 64 years with a mean of $M = 36.90$ ($SD = 5.74$).

3.2 | Parent–teacher informant agreement using the SDQ

From the GLIMMIX model, we found significant differences in SDQ scores by country in both parent and teacher reports. Table 2 shows the mean and standard deviation of the total difficulties scores and the five SDQ subscales. Figure 1 shows the distribution of externalizing and internalizing caseness by country and informant. Teachers in general reported more externalizing caseness and less internalizing caseness than parents. Lithuania reported the highest rates of externalizing caseness (10.89% and 9.8% for parent and teacher, respectively), whereas mothers in Turkey (7.14%) and teachers in Romania (4.61%) reported the highest rates of internalizing caseness.

Table 3 presents the cross-informant correlation by country. The full sample correlation varied between 24 (Prosocial) and .48 (Hyperactivity) for the five subscales. The correlation for the total difficulties scores in parents and teachers ranged from .27 (Romania) to .61 (Italy) for the seven countries, with an omnicultural mean (average of the seven countries) of 0.43. Kappa coefficients were conducted to assess

TABLE 1 Descriptive data of the full study population

Variable	N	%
Sex		
Female	2,431	49.68
Male	2,462	50.32
Age		
≤8	2,219	46.48
>8	2,555	53.52
Mother age		
<36	2,223	42.50
36–40	1,582	30.24
>40	1,426	27.26
Mother education		
>High School	2,125	49.58
Secondary completed	1,590	37.1
Some Primary or secondary	571	13.32
Mothers employment		
Employed	3,654	91.24
Unemployed	351	8.76
Number of Children		
1 child	1,383	28.26
2–3 children	2,892	59.09
More than 3 children	619	12.65
Single mothers		
2 "parent" household	3,998	85.03
1 "parent" household	704	14.97
Eastern vs. Western		
East	3,497	71.45
West	1,397	28.55

categorical agreement of externalizing and internalizing caseness (problem behavior with impact) between parent and teacher. Table 4 presents the kappas and their 95% confidence intervals for interinformant agreement for internalizing and externalizing caseness by country. Parent–teacher agreement ranged from .01 (Turkey) to .44 (Italy) for internalizing caseness and .19 (Romania) to .44 (Italy) for externalizing caseness. The confidence intervals of the kappa coefficients for internalizing caseness was .08–.18 and for externalizing caseness was .25–.35. These confidence intervals do not overlap indicating a difference between internalizing and externalizing caseness agreement. That is to say, informants agree more on externalizing problems than internalizing. Among the 346 case that met the criteria for internalizing caseness by parent or teacher report, the two informants agreed 8.96% of the time; and among the 567 case with externalizing caseness, parents and teachers agreed 20.81% of the time.

Using the chi-square to examine country differences in parent–teacher level of agreement on the caseness we found that they were significant for both internalizing caseness and externalizing caseness. Specifically, using the column proportion test, we found that for externalizing caseness, Lithuanian parent–teacher dyad was less likely to agree than parent–teacher dyad in Italy and Netherlands, and for internalizing caseness, we found that parent–teacher dyad in Italy was more likely to agree than parent–teacher dyads in other countries.

3.3 | Factors that may influence agreement

Table 5 presents the unadjusted and adjusted analyses of the associations between the predictors (child and family characteristics, parenting dimensions, and maternal distress) and parent–teacher agreement.

3.3.1 | Child, maternal, and family characteristics

Some of the risk factors that increased the likelihood of parent–teacher discrepancies were single-parent household and number of children in the household. Child older age also increased the likelihood of parent–teacher discrepancies in internalizing caseness, and male child, mother's unemployment status, and younger age also increased the likelihood of discrepancies in externalizing caseness. With the exception of gender and number of children for externalizing caseness, most variables that predicted discrepancies in the unadjusted analyses were no longer significant in the adjusted analyses.

3.3.2 | Region

In the unadjusted analysis, Western European region status was associated with less discrepancy, ORs 0.57 and 0.64 for internalizing and externalizing caseness, respectively. However, after adjusting for other significant predictors, European region was no longer significant, indicating the relationship might be explained by child, maternal, family characteristics, and parenting dimensions.

TABLE 2 Mean and standard deviation for the Strengths and Difficulties Questionnaire across seven countries

		Bulgaria		Germany		Italy		Lithuania		Netherland		Romania		Turkey	
		Mean	*SD*	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*
Total Difficulties	Parent	10.07	5.37	8.99	5.78	5.78	4.84	11.73	5.51	7.13	5.58	10.25	5.52	10.89	5.42
	Teacher	7.87	7.02	7.34	6.04	5.71	5.79	9.39	6.26	6.23	5.66	8.93	6.51	8.78	5.94
Emotional	Parent	2.19	1.94	1.97	1.92	1.53	1.70	2.96	2.21	1.98	2.14	2.64	2.22	2.55	2.23
	Teacher	1.68	2.05	1.39	1.83	1.50	1.92	2.18	2.14	1.37	2.00	2.30	2.20	2.27	2.12
Conduct	Parent	1.68	1.54	1.86	1.63	0.96	1.36	2.15	1.62	1.06	1.44	1.65	1.62	1.62	1.45
	Teacher	1.41	2.06	1.37	1.78	0.92	1.69	1.59	1.98	0.84	1.48	1.58	2.06	1.23	1.65
Hyperactivity	Parent	3.97	2.22	3.43	2.42	2.46	2.24	4.44	2.27	2.89	2.51	3.55	2.22	4.08	2.56
	Teacher	2.89	2.90	2.81	2.51	2.29	2.48	3.56	2.90	2.61	2.89	2.98	2.73	3.05	2.63
Peer Problems	Parent	2.27	1.65	1.74	1.98	0.83	1.26	2.20	1.70	1.21	1.73	2.39	1.59	2.66	1.54
	Teacher	1.90	1.80	1.81	2.06	1.00	1.54	2.06	1.83	1.41	1.87	2.08	1.61	2.25	1.63
Prosocial	Parent	7.69	1.78	8.16	1.75	8.56	1.48	7.58	1.75	8.54	1.65	8.49	1.59	8.06	1.86
	Teacher	7.58	2.44	7.66	2.36	8.16	2.12	7.16	2.38	7.90	2.39	7.95	2.23	8.06	2.02

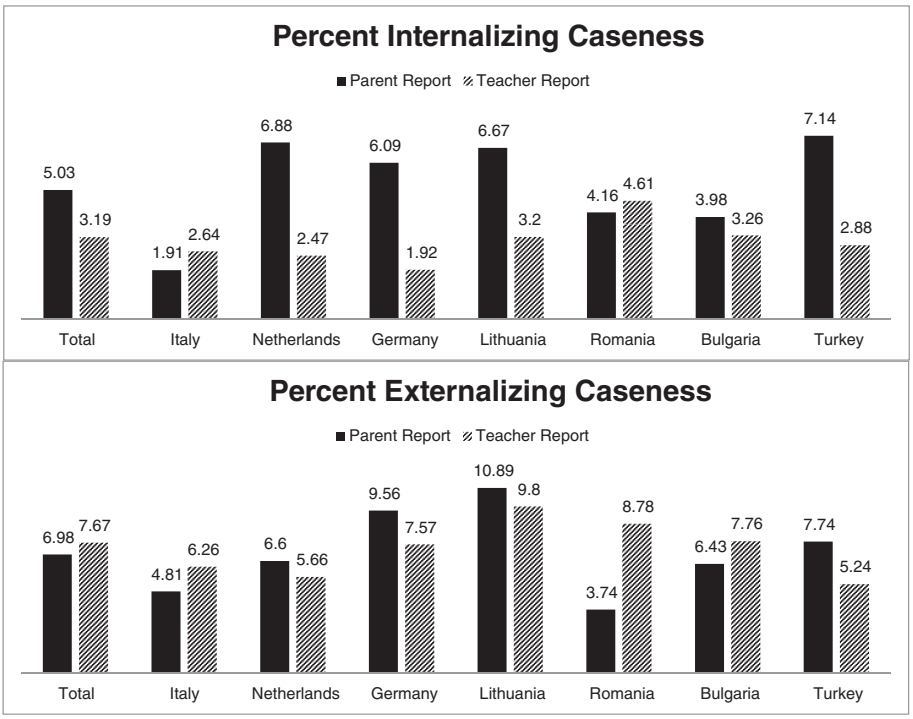

FIGURE 1 Percent parent and teacher internalizing and externalizing caseness by country

TABLE 3 Parent–teacher correlations on the Strengths and Difficulties Questionnaire

Correlation	Total difficulties	Conduct	Hyperactivity	Peer problems	Emotional	Prosocial
Italy (*N* = 687)	.61	.52	.64	.49	.48	.39
Netherlands (*N* = 531)	.53	.39	.53	.48	.47	.32
Germany (*N* = 370)	.55	.42	.58	.52	.33	.22
Lithuania (*N* = 995)	.38	.37	.50	.29	.16	.25
Romania (*N* = 932)	.27	.34	.34	.07	.15	.15
Bulgaria (*N* = 911)	.38	.38	.41	.26	.15	.16
Turkey (*N* = 468)	.29	.30	.35	.14	.16	.10
Total (*N* = 4,894)	.44	.40	.48	.35	.26	.24
Mean[a]	.43	.39	.48	.32	.27	.23

[a]Omnicultural mean for each subscale (mean of the seven correlations).

TABLE 4 Parent–teacher kappa coefficient on the Strengths and Difficulties Questionnaire internalizing caseness and externalizing caseness

Country	Internalizing with impact Kappa coefficient (95% confidence interval)	Externalizing with impact Kappa coefficient (95% confidence interval)
Italy	.44 (.22–.66)	.44 (.30–.59)
Netherlands	.30 (.13–.47)	.36 (.20–.52)
Germany	.11 (−.06–.29)	.37 (.20–.53)
Lithuania	.03 (−.05–.10)	.30 (.21–.40)
Romania	.12 (.00–.23)	.19 (.09–.30)
Bulgaria	.10 (−.02–.21)	.22 (.11–0.33)
Turkey	.01 (−.09–.11)	.24 (.08–.40)
Total	.13 (0.08–.18)	.30 (.25–.34)

3.3.3 | Parenting dimension and maternal distress

Some of the parenting dimensions increased the likelihood for parent–teacher discrepancies; these dimension include mother's low caring behavior, laxness (permissive discipline), and overreactivity (displays of anger, meanness and irritability). The relationship between a parent–teacher disagreement and mother's low caring attitude and laxness seized to exist when considering other demographic characteristics, region, and parenting dimensions. Overreactivity (ORs, 2.65 and 2.18 for internalizing and externalizing caseness, respectively) remained the only robust predictor in the multiple regression analysis. Maternal distress and mother's autonomy-promoting attitudes were not related to parent–teacher disagreement in neither internalizing nor externalizing caseness.

TABLE 5 Association between child and family characteristics, country region, parenting styles, and maternal distress factors on disagreement in internalizing and externalizing caseness between parent and teacher

	Unadjusted						Adjusted					
	Internalizing with impact			Externalizing with impact			Internalizing with impact			Externalizing with impact		
	OR	Lower 95% CI	Upper 95% CI	OR	Lower 95% CI	Upper 95% CI	OR	Lower 95% CI	Upper 95% CI	OR	Lower 95% CI	Upper 95% CI
Male	1.31	1.04	1.65	2.13*	1.73	2.62				2.39*	1.80	3.18
Child age ≤ 8 vs. <8	1.18*	1.08	1.29	1.02	0.95	1.10	1.16	1.03	1.31			
Mother age (continuous)	0.98	0.96	1.00	0.97*	0.95	0.99				0.97	0.95	1.00
Maternal education												
Secondary Completed vs. <high school	1.50	1.16	1.95	1.32	1.06	1.66						
Some primary or secondary vs. <high school	1.34	.91	1.97	1.50	1.10	2.04						
Mothers unemployment	1.74	1.18	2.56	1.72*	1.24	2.40				1.31	0.85	2.01
Number of children	1.15*	1.07	1.25	1.17*	1.10	1.25	1.13	1.01	1.27	1.17*	1.06	1.29
Single parent	1.86*	1.40	2.47	1.66*	1.30	2.13	1.71	1.09	2.69	1.23	0.82	1.84
Western Europe	0.57*	0.43	0.76	0.64*	.51	0.81	1.12	0.73	1.70	0.91	0.62	1.33
Parenting dimensions												
Low caring	1.78*	1.35	2.34	1.96*	1.55	2.47	1.26	0.85	1.87	1.28	0.93	1.79
Low autonomy	0.84	0.61	1.16	0.91	0.69	1.19						
High laxness	1.79*	1.31	2.44	1.55	1.18	2.04	1.40	0.91	2.16			
High overreactivity	2.87*	2.17	3.80	2.70*	2.11	3.44	2.65*	1.84	3.81	2.18*	1.59	2.99
High verbosity	1.13	0.80	1.59	1.05	0.78	1.41						
High maternal distress	4.16	0.57	30.15	1.59	0.58	4.39						

Note. CI = confidence interval; OR = odds ratio.

*$p \leq .001$.

4 | DISCUSSION

This study used a sample population of more than 4,000 children in seven diverse European countries to (a) examine whether parent–teacher agreements using the SDQ varied by problem behaviors and by countries and (b) examine explanatory factors for informant disagreement such as child, maternal, and family characteristics, parenting dimensions, and maternal distress.

4.1 | Parent–teacher informant agreement using the SDQ

As hypothesized, we found that externalizing behaviors tended to have greater levels of cross-informant agreement than internalizing behaviors. This is because externalizing behaviors are more overt and observable than internalizing behaviors. Analogous to Rescorala et al. (2014) study, we also found that cross country differences on parent–teacher agreement on externalizing and internalizing behavior varied significantly in the seven European countries.

Overall, there were several findings consistent with previous review (Achenbach et al., 1987; Rescorla et al., 2014; Stone et al., 2010). First, parent–teacher agreements were low to modest. Discrepancies were found when using both Pearson correlation and kappa coefficient. The parent–teacher correlation for the total difficulties score (.44) mirrored those reported by Stone et al. (2010). The overall parent–teacher correlation reported by the SDQ was higher than the overall parent–teacher correlation reported by the meta-analyses of both Achenbach et al. (1987) and De Los Reyes et al. (2015), which was r = .28. Parent–teacher agreement using kappa was .13 (internalizing caseness) and .30 (externalizing caseness) and is comparable to those reported by Rescorla et al. (2014), which were .15 and .21 for internalizing and externalizing behavior, respectively. However, when examining country specific kappa, several differences emerged. For example, the kappa's reported by Rescorla et al. (2014) were higher than in this study for several countries, including in Lithuania (kappa = .21 for internalizing), Romania (kappa = .24 for internalizing), and Turkey (kappa = .13 for internalizing). These differences are likely a function of the measures used. Rescorla et al. (2014) examined the CBCL and TRF, while this study examined the SDQ, suggesting that parent–teacher agreement can vary by instruments used.

Although parent–teacher agreement was low to moderate in this study, it is important to remember that the discrepancies do not indicate a lack of valid judgment of one informant over another but can be a result of unique information provided by each (Van Roy et al., 2010). For example, parent and teacher observe the child in different settings and can have different thresholds for the same problem behaviors in different situations (Achenbach et al. 1987; De Los Reyes et al., 2015; Hartley, Zakriski, & Wright, 2011). It has also been hypothesized that children with externalizing and/or internalizing problems might be characterized differently depending on whether the problem is endorsed by parent only, teacher only, or by both informants. Acknowledging these differences and documenting the effects on diagnostic classifications across different countries are vital to further examine the implications of informant discrepancies for children with internalizing and/or externalizing behavior (Achenbach, 2005; Carlson & Dyson, 2011; De Los Reyes & Kazdin, 2004).

4.2 | Factors that may influence agreement

We hypothesized that parent–teacher disagreement will be varied by child age, maternal education status, age and employment status, family structure, parenting dimension, and maternal stress. In fact, we found that some child and family characteristics and different parenting dimensions increased the likelihood of parent–teacher disagreement. Consistent with our hypothesis, we found single status to be associated with more disagreement. This might be because single parenthood result in less cohesiveness and control leading to lower parent–teacher agreement. Single parenthood might also be associated with limited access to social and financial support, and in turn associated with more parental stress, which has been shown to be associated with parent–teacher agreement (Van der Oord et al., 2006). Furthermore, conditions in home and school might be more different, leading children to behave differently at home and at school than is the case with children from better regulated households. Number of children in household was also related to more discrepancies. This could be because attention and affection have to be shared among multiple children (Michels et al., 2013). We found child gender to significantly influence parent–teacher agreement in externalizing problem but not internalizing, suggesting the need to further explored child gender as a moderator. Child age was found only to be significantly associated with internalizing caseness and not to externalizing caseness. This might provide insight as to why past studies have mixed findings, suggesting the need to look at child age with problem behavior type. Also contrary to the literature, our study did not find maternal distress to be associated with more disagreement (Briggs-Gowan et al., 1996; De Los Reyes & Kazdin, 2006; Ehrlich et al., 2013; Richters, 1992). However, this was consistent with Van der Oord et al. (2006) study, which found that it was parenting stress and not parental depression that influenced parent-agreement on attention deficit hyperactivity disorder and oppositional defiant disorder.

Few studies in the past have examined parenting dimension as an explanatory factor for informant discrepancies although it has been hypothesized that negative parenting practice might be associated with more discrepancies (De Los Reyes et al., 2013). One known study that examined parenting dimension and informant discrepancies was Michels et al. (2013) study. They found that authoritative parenting was associated with less parent–teacher disagreement. The results of this study further support the findings that parenting dimensions are significant predictors of parent–teacher disagreement, specifically that dysfunctional parenting discipline such as laxness and overreactivity (displays of anger, meanness, and irritability) were related to more informant discrepancies. Mothers with high overreactivity might have a lower threshold for child problem behaviors (i.e., externalizing and internalizing behavior) and children of mother with high overreactivity might be less likely to share emotions with their parents, therefore leading to more discrepancies. Mother's low-caring attitude was also related to more disagreement. Less-caring parents might be less in tune with their children's feelings and problem behaviors or lack of, therefore leading to more parent–teacher discrepancies. To the best

of our knowledge, this was one of the first studies to examine multiple dimensions of parenting dimensions.

4.2.1 | Limitation

Despite the strengths of this study, there were several limitations. Although internal consistency of the total difficulties scales was satisfactory, selected subscales of the SDQ were less reliable. However, this mirrors the results of Stone et al. (2010). Goodman, Lamping, and Ploubidis (2010) suggested that using the broader internalizing and externalizing SDQ subscales for analyses in low-risk samples might be more suitable. In fact, we found that reliability for hyperactivity and conduct problems combined performed much better for parents, α = .73 (Lithuania)–.82 (Italy), and teachers, α = .82 (Turkey)–.89 (Bulgaria). This is also one of the reasons why we decided to examine externalizing and internalizing caseness rather than the five subscales of SDQ. Another limitation was that we restricted our sample to mothers only. Including different caregivers (e.g., fathers and grandparents) might yield different results. Parenting dimensions were based on maternal self-reports and may be a biased view of their parenting practices. Parenting practice is known to be influenced by other characteristics (e.g., maternal psychopathology, child behavior, and context) and measuring that in different ways might provide different results. Future research should consider other informants and the use of other methods (e.g., observations of parenting instead of self-report) to collect information on child behavior, parenting construct, and parent psychopathology.

In this study, we used correlation and kappa to test for informant's relative agreement. Possible explanatory factors for informant discrepancies were explored using logistic regression as a way to understand underlying mechanisms and factors that play important roles in informant discrepancies. Although correlation has been one of the most common approach to assess informant discrepancies, in the last decade, different ways to study informant discrepancies have emerged, all with their specific statistical and conceptual pros and cons (Laird & De Los Reyes, 2013). Studies in adults have used growth models to address concerns regarding measurement errors and future studies might want to consider this approach (Harvey et al., 2013; Laird & De Los Reyes, 2013; Laird & Weems, 2011).

5 | CONCLUSION

Despite these limitations, this study adds to our knowledge regarding multi-informant discrepancies using the SDQ in seven different European countries, including both Eastern and Western Europe. This study also advances our understanding regarding the factors that play important roles in influencing discrepancies. This is also one of the few studies that explored the influences of parenting dimension on informant discrepancies. Furthermore, the different factors that relate to informant discrepancies remind us that child mental health outcome does not exist independent of parenting practices, child characteristics (e.g., gender and age) or other aspects of the child's life (e.g., country of residence and family life). Understanding the factors that contribute to the discrepancies of children's behavior problems has potential implications for diagnostic assessment of behavior disorder in school children.

ORCID

Viviane Kovess-Masfety http://orcid.org/0000-0001-7045-6175

REFERENCES

Aaronson, N. K., Acquadro, C., Alonso, J., Apolone, G., Bucquet, D., Bullinger, M., … & Razavi, D. (1992). International quality of life assessment (IQOLA) project. *Quality of Life Research*, *1*(5), 349–351.

Achenbach, T. M., McConaughy, S. H., & Howell, C. T. (1987). Child/adolescent behavioral and emotional problems: Implications of cross-informant correlations for situational specificity. *Psychological Bulletin*, *101*(2), 213–232.

Achenbach, T. M. (2005). Advancing assessment of children and adolescents: Commentary on evidence-based assessment of child and adolescent disorders. *Journal of Clinical Child and Adolescent Psychology*, *34*(3), 541–547.

Arnold, D. S., O'Leary, S. G., Wolff, L. S., & Acker, M. M. (1993). The Parenting Scale: A measure of dysfunctional parenting in discipline situations. *Psychological Assessment*, *5*(2), 137.

Bergeron, L., Valla, J.-P., & Breton, J.-J. (1992). Pilot study for the Quebec Child Mental Health Survey: Part II. Correlates of DSM-III-R criteria among six to 14 year olds. *Canadian journal of psychiatry Revue canadienne de psychiatrie*, *37*(6), 381–386.

Briggs-Gowan, M. J., Carter, A. S., & Schwab-Stone, M. (1996). Discrepancies among mother, child, and teacher reports: Examining the contributions of maternal depression and anxiety. *Journal of Abnormal Child Psychology*, *24*(6), 749–765.

Carlson, G. A., & Dyson, M. (2011). Diagnostic implications of informant disagreement about rage outbursts: Bipolar disorder or another condition? *The Israel Journal of Psychiatry and Related Sciences*, *49*(1), 44–51.

Chi, T., & Hinshaw, S. (2002). Mother–child relationships of children with ADHD: The role of maternal depressive symptoms and depression-related distortions. *Journal of Abnormal Child Psychology*, *30*(4), 387–400. https://doi.org/10.1023/A:1015770025043

De Los Reyes, A. (2011). Introduction to the special section: More than measurement error: Discovering meaning behind informant discrepancies in clinical assessments of children and adolescents. *Journal of Clinical Child & Adolescent Psychology*, *40*(1), 1–9. https://doi.org/10.1080/15374416.2011.533405

De Los Reyes, A., & Kazdin, A. E. (2004). Measuring informant discrepancies in clinical child research. *Psychological Assessment*, *16*(3), 330.

De Los Reyes, A., & Kazdin, A. E. (2005). Informant discrepancies in the assessment of childhood psychopathology: A critical review, theoretical framework, and recommendations for further study. *Psychological Bulletin*, *131*(4), 483.

De Los Reyes, A., & Kazdin, A. E. (2006). Informant discrepancies in assessing child dysfunction relate to dysfunction within mother–child interactions. *Journal of Child and Family Studies*, *15*(5), 643–661.

De Los Reyes, A., Thomas, S. A., Goodman, K. L., & Kundey, S. M. (2013). Principles underlying the use of multiple informants' reports. *Annual Review of Clinical Psychology*, *9*, 123–149.

De Los Reyes, A., et al. (2015). The validity of the multi-informant approach to assessing child and adolescent mental health. *Psychological Bulletin*, *141*(4), 858–900. https://doi.org/10.1037/a0038498

Duhig, A. M., Renk, K., Epstein, M. K., & Phares, V. (2000). Interparental agreement on internalizing, externalizing, and total behavior problems: A meta-analysis. *Clinical Psychology: Science and Practice*, *7*(4), 435–453.

Ehrlich, K. B., Cassidy, J., Lejuez, C. W., & Daughters, S. B. (2013). Discrepancies about adolescent relationships as a function of informant attachment and depressive symptoms. *Journal of Research on Adolescence:n/a-n/a*. https://doi.org/10.1111/jora.12057

Fung, J. J., & Lau, A. S. (2010). Factors associated with parent–child (dis) agreement on child behavior and parenting problems in Chinese immigrant families. *Journal of Clinical Child & Adolescent Psychology*, *39*(3), 314–327.

Goodman, R. (1997). The Strengths and Difficulties Questionnaire: A research note. *Journal of Child Psychology and Psychiatry*, *38*(5), 581–586.

Goodman, R. (1999). The extended version of the Strengths and Difficulties Questionnaire as a guide to child psychiatric caseness and consequent burden. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *40*(05), 791–799.

Goodman, A., Lamping, D., & Ploubidis, G. (2010). When to use broader internalising and externalising subscales instead of the hypothesised five subscales on the Strengths and Difficulties Questionnaire (SDQ): Data from British parents, teachers and children. *Journal of Abnormal Child Psychology*, *38*(8), 1179–1191. https://doi.org/10.1007/s10802-010-9434-x

Guion, K., Mrug, S., & Windle, M. (2009). Predictive value of informant discrepancies in reports of parenting: Relations to early adolescents' adjustment. *Journal of Abnormal Child Psychology*, *37*(1), 17–30.

Hartley, A. G., Zakriski, A. L., & Wright, J. C. (2011). Probing the depths of informant discrepancies: Contextual influences on divergence and convergence. *Journal of Clinical Child & Adolescent Psychology*, *40*(1), 54–66.

Harvey, E. A., Fischer, C., Weieneth, J. L., Hurwitz, S. D., & Sayer, A. G. (2013). *Predictors of discrepancies between informants' ratings of preschool-aged children's behavior: An examination of ethnicity, child characteristics, and family functioning*. Early Childhood Research Quarterly.

Jensen, P. S., Xenakis, S. N., Davis, H., & Degroot, J. (1988). Child psychopathology rating scales and interrater agreement: II. Child and family characteristics. *Journal of the American Academy of Child & Adolescent Psychiatry*, *27*(4), 451–461.

Kovess, V., Carta, M. G., Pez, O., Bitfoi, A., Koç, C., Goelitz, D., … Otten, R. (2015). *The School Children Mental Health in Europe (SCMHE) Project: Design and first results. Clinical practice and epidemiology in mental health: CP & EMH*, 11(Suppl 1 M7), 113–123.

Laird, R., & De Los Reyes, A. (2013). Testing informant discrepancies as predictors of early adolescent psychopathology: Why difference scores cannot tell you what you want to know and how polynomial regression may. *Journal of Abnormal Child Psychology*, *41*(1), 1–14. https://doi.org/10.1007/s10802-012-9659-y

Laird, R. D., & Weems, C. F. (2011). The equivalence of regression models using difference scores and models using separate scores for each informant: Implications for the study of informant discrepancies. *Psychological Assessment*, *23*(2), 388–397. https://doi.org/10.1037/a0021926

Michels, N., et al. (2013). Parental and children's report of emotional problems: Agreement, explanatory factors and event-emotion correlation. *Child and Adolescent Mental Health*, *18*(3), 180–186.

Pelton, J., Steele, R., Chance, M., & Forehand, R. (2001). Discrepancy between mother and child perceptions of their relationship: II. Consequences for children considered within the context of maternal physical illness. *Journal of Family Violence*, *16*(1), 17–35. https://doi.org/10.1023/A:1026572325078

Pez, O., Boyd, A., Fermanian, C., & Kovess-Masfety, V. (2011). *The School Children Mental Health project final report 2011 DG Health and consumers (SANCO) EU*.

Rescorla, L. A., et al. (2013). Cross-informant agreement between parent-reported and adolescent self-reported problems in 25 societies. *Journal of Clinical Child & Adolescent Psychology*, *42*(2), 262–273.

Rescorla, L. A., et al. (2014). Parent–teacher agreement on children's problems in 21 societies. *Journal of Clinical Child & Adolescent Psychology*, *43*(4), 627–642. https://doi.org/10.1080/15374416.2014.900719

Richters, J. E. (1992). Depressed mothers as informants about their children: A critical review of the evidence for distortion. *Psychological Bulletin*, *112*(3), 485.

Stone, L. L., Otten, R., Engels, R. C. M. E., Vermulst, A. A., & Janssens, J. M. A. M. (2010). Psychometric properties of the parent and teacher versions of the Strengths and Difficulties Questionnaire for 4- to 12-year-olds: A review. *Clinical Child and Family Psychology Review*, *13*(3), 254–274. https://doi.org/10.1007/s10567-010-0071-2

Stone, S., Speltz, M., Collett, B., & Werler, M. (2013). Socioeconomic factors in relation to discrepancy in parent versus teacher ratings of child behavior. *Journal of Psychopathology and Behavioral Assessment*, *35*(3), 314–320. https://doi.org/10.1007/s10862-013-9348-3

Treutler, C. M., & Epkins, C. C. (2003). Are discrepancies among child, mother, and father reports on children's behavior related to parents' psychological symptoms and aspects of parent–child relationships? *Journal of Abnormal Child Psychology*, *31*(1), 13–27.

Van der Ende, J., & Verhulst, F. C. (2005). Informant, gender and age differences in ratings of adolescent problem behaviour. *European Child & Adolescent Psychiatry*, *14*(3), 117–126.

Van der Oord, S., Prins, P. J., Oosterlaan, J., & Emmelkamp, P. M. (2006). The association between parenting stress, depressed mood and informant agreement in ADHD and ODD. *Behaviour Research and Therapy*, *44*(11), 1585–1595.

Van Roy, B., Groholt, B., Heyerdahl, S., & Clench-Aas, J. (2010). Understanding discrepancies in parent-child reporting of emotional and behavioural problems: Effects of relational and socio-demographic factors. *BMC Psychiatry*, *10*(1), 56.

Viera, A. J., & Garrett, J. M. (2005). Understanding interobserver agreement: The kappa statistic. *Family Medicine*, *37*(5), 360–363.

Wadsworth, S. J., Corley, R. P., Hewitt, J. K., & DeFries, J. C. (2001). Stability of genetic and environmental influences on reading performance at 7, 12, and 16 years of age in the Colorado Adoption Project. *Behavior Genetics*, *31*(4), 353–359. https://doi.org/10.1023/A:1012218301437

How to cite this article: Cheng S, Keyes KM, Bitfoi A, et al. Understanding parent–teacher agreement of the Strengths and Difficulties Questionnaire (SDQ): Comparison across seven European countries. *Int J Methods Psychiatr Res*. 2018;27: e1589. https://doi.org/10.1002/mpr.1589