

Andreas Becker
Nicola Hagenberg
Veit Roessner
Wolfgang Woerner
Aribert Rothenberger

Evaluation of the self-reported SDQ in a clinical setting: Do self-reports tell us more than ratings by adult informants?

■ **Abstract** *Objective* The aim of this study was to evaluate the German self-reported Strengths and Difficulties Questionnaire (SDQ) in a clinical setting. We also investigated whether this additional information gathered directly from older children and adolescents improves the prediction of clinical status when external ratings from their parents and/or teachers are already available. *Methods* SDQ

self-reports were collected from 214 in- and outpatients (81 girls and 133 boys) aged 11 to 17 years who were seen at the department of child and adolescent psychiatry of the University of Göttingen. Results obtained with the self-rated questionnaire were compared with the parent and teacher SDQs, corresponding CBCL/YSR scores, and the clinical diagnostic classification. Finally, the additional diagnostic benefits of the self-reports were examined. *Results* The scales of the SDQ self-report proved to be sufficiently homogeneous, and acceptable correlations were found with the equivalent parent and teacher ratings. The self-rated version of the SDQ demonstrated good validity with respect to the differentiation between clinically

defined cases and non-cases and in detecting various subcategories of psychiatric disorders within the clinic sample. SDQ self-reports significantly contributed to the prediction of diagnostic status, specifically if only parent or teacher ratings were available. *Conclusions* The self-rated version of the SDQ was shown to be a reliable and valid method for the assessment of behavioural problems in children and adolescents. In the absence of adult informant reports from parents and teachers, the diagnostic value of self-ratings was also demonstrated.

■ **Key words** Strengths and Difficulties Questionnaire (SDQ) – self-report – YSR – ROC analysis

A. Becker (✉) · N. Hagenberg · V. Roessner · W. Woerner · A. Rothenberger
University of Göttingen
Child and Adolescent Psychiatry
von-Siebold-Str. 5
37075 Göttingen, Germany
Tel.: +49-551/39-2386
Fax: +49-551/39-2696
E-Mail: abecker4@gwdg.de

Introduction

Various studies have shown that a considerable percentage of children and adolescents meet the criteria for emotional or behavioural disorders [9, 26]. Although their problems can create much distress for daily life and well-being, many adolescents with psychiatric disorders receive no professional help, even though early intervention could prevent persistence of the disorder into adult life [16].

Evaluation of child psychiatric disturbances is mainly based on clinical interviews with parents and teachers, assessment of problem behaviour with various questionnaires, and on observations of behaviour in a

diagnostic setting. For a comprehensive evaluation of such disturbances it is necessary to draw on information from the child or adolescent him/herself as a valuable source for describing the patient's feelings, moods, and subjective experiences [32].

For this purpose, structured interviews such as the Diagnostic Interview of Psychological Disorders in children and adolescents (German child DIPS [29]), the Kiddie-SADS interview [18] or the NIMH Diagnostic Interview Schedule for Children (DISC [7]) can be employed. In these assessment instruments, the interview questions are standardised to reduce variance of both applied criteria and collected information, thus increasing reliability and allowing a better comparison between patients or studies. However, these methods are mainly

applied in the framework of scientific investigations. Due to the considerable amount of time required as well as the necessary degree of clinical expertise of the interviewer, structured interviews are of limited importance for primary health care in a clinical setting [17]. Given these problems and limitations, child and adolescent psychiatry has benefited from the development and application of screening questionnaires which are shorter, less costly, and easier to apply [10]. Using such tools for gathering information, the clinician can quickly obtain an impression of the most relevant psychopathological symptoms in children or adolescents.

Since younger children only have a limited ability for introspection and are often unable to judge and report on their emotions or behaviour, their direct assessment is considered to be only of limited diagnostic value [3]. Children under the age of 11 years are also often unable to transfer their daily experiences onto a questionnaire made up of only a few items. Thus, it seems reasonable to only utilise the information from self-reports of children aged about 11 years or older.

Several studies have shown that the act of seeking clinical assistance has considerable influence on the validity of the characteristics being assessed. For example, a distinct answer tendency in a non-clinical random sample showed that children and adolescents described themselves as having more behavioural problems than were reported by their parents and teachers [27]. The opposite effect was demonstrated in a clinical sample [14], where dissimulation tendencies made children and adolescents under-report their behavioural problems.

Methods and results when examining the utility of self-reports and their contribution to clinical diagnostics were found to be very heterogeneous. One approach to assess the quality of the information obtained with the self-rated version of the SDQ involves comparisons with the well-evaluated Youth Self-Report (YSR; [2]). The YSR is the self-report derivative of the Child Behavior Checklist (CBCL; [1]), and has been widely implemented to describe relevant problems in children and adolescents [6]. Initial reports on the internal consistency of composite scales of the YSR [2] were later replicated in Switzerland [28], Norway [22], and Germany [25]. Cross-informant agreement between CBCL parent reports and YSR self-ratings usually yielded only moderate correlations (e. g., [19, 30]). Nevertheless, Morgan and coworkers [23] successfully predicted clinical (DSM-III R) diagnosis on the basis of YSR subscales.

However, it must be noted that the scale structure of the YSR was empirically determined, and only some of its problem scales correspond to specific diagnostic categories. For example, although several symptoms of attention-deficit and hyperactivity disorders are included in the attention problems subscale of the YSR, the items on this scale do not cover the entire range of criteria specified for the corresponding clinical definitions ac-

ording to DSM-IV or ICD-10. As a consequence, the validity of this YSR subscale assessing attention disorders was only partly confirmed in a clinical sample [8].

On the other hand, the Strengths and Difficulties Questionnaire (SDQ) was designed by Goodman [11] on the basis of nosological criteria (ICD-10). Like its parent- and teacher-rated counterparts, the self-report version of the SDQ only comprises 25 items and addresses a well-balanced number of positive and negative behavioural aspects, in contrast to the considerably longer YSR. In Germany, normative data have only been established for the parent-rated version of the SDQ ([31]; see also Woerner et al., in this supplement volume).

Although the SDQ has been used as a diagnostic instrument only since 1997, studies from several countries have already reported psychometric properties of its self-rated version. Results on the internal consistency (Cronbach's α) demonstrate the homogeneous scale structure, with reliabilities for the total difficulties score ranging between 0.82 (England: [13]), 0.78 (Netherlands [24]), and 0.71 (Finland: [20]). In spite of their shortness, good respective results were also obtained for the five self-rated subscales.

Analyses of cross-informant agreement yielded a similar correspondence between the self-rated SDQ and the parent version as had been found for the YSR: Correlations for the total difficulties score of the SDQ obtained in clinical or community-based sample in the Netherlands, in Finland, and in the UK ranged from 0.40 to 0.46.

A review of studies with structured diagnostic interviews [17] demonstrated that the level of agreement between self-reports and parent ratings depends on the type of disorder. Similar findings were noted with the SDQ in a clinical random sample [13], with a higher degree of correspondence between parent and self-rating for externalising disorders and lower cross-informant agreement for internalising disorders. However, the opposite tendency was observed in community samples [12, 24].

Heterogeneous results were also reported with respect to mean scale scores of adult informants and self-reports. In an English clinical sample, Goodman and colleagues [13] found that older children and adolescents often rated themselves in the normal or borderline range of the total difficulties score, while their parents or teachers usually reported a higher degree of impairment. In contrast, Finnish children and adolescents described themselves as having considerably more behavioural problems than their parents did [20].

Consistent gender differences in self-rated SDQ scores were documented in a large number of reports. For example, Dutch girls had significantly higher values on the emotional symptoms and prosocial behaviour subscale, while boys scored higher on the conduct problems scale [24]. Similar findings were reported from Fin-

land [21], where girls in their 7th or 9th school year scored much higher than boys on emotional symptoms and prosocial behaviour, but also had a higher total difficulties score. In this Finnish sample, boys again reported more conduct problems than girls.

To evaluate the validity of the self-rated SDQ, it was compared with the corresponding YSR scores. In Finland, this yielded correlations of 0.71 for the total difficulties score and between 0.43 and 0.68 for the subscales. In a German non-clinical random sample, Bettge and coworkers [5] obtained a somewhat lower correlation of 0.58 for the total difficulties score.

Based on these previous results gathered for the self-rated SDQ, the present study comprehensively evaluates the German version in a clinical setting. We examined how well the information from children and adolescents matches with the judgements by their parents and teachers. To further analyse the psychometric properties of the SDQ self-report scales, these were compared with the corresponding YSR scores and with the clinical diagnosis as determined according to ICD-10 criteria. Here, it was examined how accurately the self-report version of the SDQ is able to discriminate between older children with or without any diagnosis of a child psychiatric disorder. At a more detailed level, SDQ scores were used to detect more specific diagnostic subgroups. Finally, another question of particular interest was whether the self-ratings gathered from children and adolescents additionally improve the prediction of their clinical diagnosis when external ratings from parents and/or teachers are also available.

Methods

Clinical sample and diagnostic subgroups

In the present study, data were collected from in- and outpatients of the Child and Adolescent Psychiatry Department of the University of Göttingen and from their parents and teachers. At the time of their assessment (August 1998 until July 2000) all patients were between 11 and 17 years old. SDQ self-reports and parent-rated SDQs were available for all participants, as well as self- and parent-rated YSR/CBCL scores. Independently determined clinical diagnoses by board-certified child and adolescent psychiatrists were based on ICD-10 criteria. The present analysis sample with complete self-report and parent data numbers 214 older children and adolescents, including 133 boys (62%) with a mean age of 13.8 years (SD: 1.9) and 81 girls (38%) with a mean age of 14.8 years (SD: 1.6).

As shown in Table 1, 194 patients (90.7% of the entire analysis sample) received at least one clinical diagnosis of a child psychiatric disorder (any diagnosis on axis I except F70 to F79, F80 to F83.99, F85 to F89, and F98).

Table 1 Clinical sample with completed self-report and parent SDQ. Distribution of diagnostic categories and mean age (N = 214)

Diagnostic categories	Boys N = 133	Girls N = 81	Total N = 214
Any diagnosis on axis I	N = 121 (91 %)	N = 73 (90.1 %)	N = 194 (90.7 %)
Emotional disorders	N = 55 (41.4 %)	N = 45 (55.6 %)	N = 100 (46.7 %)
Conduct/oppositional disorders	N = 49 (36.8 %)	N = 20 (24.7 %)	N = 69 (32.2 %)
Hyperactivity/ attention-deficit disorders	N = 40 (30.1 %)	N = 3 (3.7 %)	N = 43 (20.1 %)
Mean age (years) (SD)	13.8 (1.9)	14.8 (1.6)	14.2 (1.9)

Subdividing into broad-band diagnostic subgroups, 100 children and adolescents (46.7%) showed an emotional disorder; 69 (32.2%) had an oppositional/conduct disorder, and 43 patients (20.1%) were found in the subcategory of hyperactivity/attention-deficit disorders (see [4] for a more detailed definition of diagnostic subgroups).

In order to compare the self-report results with those obtained with parent and teacher ratings, some of the analyses were carried out for a subsample of 124 patients who had complete sets of SDQ data from all three informants (i. e. parent, teacher, and self-reports).

Statistical analysis

For all three versions of the SDQ, means and reliabilities (Cronbach's α) were calculated for the total difficulties score and each of the five subscales. Spearman rank correlations between the patients' age and SDQ scores were also determined. To examine the factor structure of the SDQ self-report, a principal component analysis with varimax rotation was carried out. Rank correlations between the self-reports and adult-rated scores were calculated for the entire analysis sample and separately for each gender subgroup.

Since normative data for the German SDQ self-report were not available at the time of writing, evaluation in this study was performed without applying predefined cut-off values. Hence, cut-off-independent Receiver Operating Characteristics (ROC) analyses were chosen to quantify and compare the discriminative validity of the SDQ. ROC analyses yield a cut-off-independent measure called AUC (area under the curve), which summarises the sensitivity and specificity of all possible cut-off positions with respect to a known external criterion (e. g. clinical diagnosis of hyperactivity/attention-deficit disorder: yes vs. no). The obtained AUC values reflect the discriminative validity, and allow direct comparisons

between the predictive value of two different instruments. An AUC of 1.0 would mean that the instrument discriminates perfectly between children with and without a certain diagnosis; an AUC of 0.5 is obtained when a measure merely discriminates at chance level. For the present study, ROC analyses were used to compare differences between both self-reports (self-rated SDQ vs. YSR), between the two parent questionnaires (parent SDQ vs. CBCL), and between self-reports and parent reports using the same instrument (self-rated vs. parent SDQ; YSR vs. CBCL). The algorithm used for statistical comparisons between two obtained AUC values has been described in detail by Hanley & McNeil [15].

Furthermore, adjusted R^2 values obtained in block-wise multiple regression analyses were examined, to show the effects of different combinations of SDQ informants on the explained variance of the dependent target variable (i. e. presence of any psychiatric disorder or a specific clinical diagnosis).

Results

■ Scale means, scale homogeneity and correlations with age

For each of the SDQ versions, Table 2 reports mean scale scores, the homogeneity coefficient of each scale, and the correlations between the patients' age and their SDQ scores.

Examination of the pattern of scale means shows that the children and adolescents in this clinical sample described themselves as having fewer problems and showing more prosocial behaviour than according to their parents' reports. In contrast, teacher ratings of their pupils mentioned fewer emotional symptoms and less prosocial behaviour.

Analysis of the reliabilities (Cronbach's α) of the SDQ

scales yielded satisfactory results (total difficulties score in self-reports: $\alpha = 0.78$; compared to $\alpha = 0.82$ for parent and teacher ratings, respectively). Although only containing five items each, subscales of the self-reported SDQ were also shown to be acceptably homogeneous (e. g. emotional symptoms: $\alpha = 0.77$). Only the internal consistency of the conduct problems subscale fell below 0.60. Subscales measuring conduct problems, hyperactivity-inattention, and peer problems were clearly more homogeneous in parent and teacher ratings than in the self-reports (Table 2).

For all three informants, correlations between the patients' age and SDQ scores showed a significant negative age effect on the conduct problems subscale (self-report: -0.19). While parents as well as teachers reported significantly fewer total difficulties and less hyperactive/inattentive behaviour with increasing age of the children, neither of these two scores was strongly related to age when obtained from the patients themselves. Conversely, older patients scored higher on self-reported emotional symptoms and prosocial behaviour than younger ones, whereas the same scores taken from parent and teacher ratings reflected either much weaker age effects or none at all.

■ Evaluation of the factorial structure

The factor structure of the self-rated SDQ within this clinic sample was examined by carrying out a principal component analysis with varimax rotation. The five extracted factors explained 51.4% of the total variance, and the resulting pattern of loadings showed a high conformance with the original SDQ scales.

Table 2 Scale evaluation of the German parent, teacher, and self-report SDQ in a clinical sample (11- to 17-year-old Göttingen in- and outpatients, N = 124 with complete data from all three informants)

SDQ scale	Scale properties								
	self-report			parent rating			teacher rating		
	scale mean	reliability ¹ α	r with age	scale mean	reliability ¹ α	r with age	scale mean	reliability ¹ α	r with age
Total Difficulties Score	15.0	0.78	0.01	17.0	0.82	-0.16*	14.7	0.82	-0.20*
Emotional Symptoms	3.8	0.77	0.31***	4.4	0.73	0.12	3.0	0.77	0.15*
Conduct Problems	3.0	0.58	-0.19*	4.0	0.79	-0.21*	2.8	0.76	-0.28***
Hyperactivity/Inattention	4.9	0.65	-0.14	5.8	0.82	-0.28***	5.3	0.87	-0.25**
Peer Problems	3.2	0.65	-0.01	3.4	0.76	-0.09	3.6	0.75	-0.08
Prosocial Behaviour	6.9	0.78	0.24**	6.0	0.76	0.07	4.5	0.82	0.18*

* $p \leq 0.05$; ** $p \leq 0.01$; *** $p \leq 0.001$ (Spearman rank correlations, one-tailed)

¹ homogeneity coefficient (Cronbach's α)

■ Evaluation of cross-informant agreement

Agreement between self-reports and adult informant-rated scores was determined by calculating rank correlations between corresponding scales (Table 3). Although all coefficients attained statistical significance, it was noted that the total difficulties score and the emotional symptoms subscale showed a much weaker cross-informant agreement than the other SDQ scores, particularly when compared to the conduct problems and peer problems subscales. Overall, self-ratings were more closely associated with their parents' reports than with those by their teachers.

When cross-informant agreement was regarded separately in male and female subsamples, correlations between self-reports and parent ratings were found to be much higher for girls than for boys. Girls' self-ratings showed significantly higher agreement with their parents' judgements in the total difficulties score (boys: 0.29; girls: 0.61) and with respect to hyperactivity/inattention, conduct problems, and peer problems. The subscales assessing emotional symptoms and prosocial behaviour did not show such gender-dependent effects,

nor did any of the correlations between self-reports and teacher ratings (Table 3).

■ Evaluation of discriminative validity

ROC analyses were carried out to evaluate the discriminative properties of the self-rated SDQ in comparison to the YSR. The predictive power of the self-report measures was also compared to parent reports (parent SDQ and CBCL). Both the ability to distinguish between patients with and without any psychiatric diagnosis and detection of diagnostic subgroups were examined in these analyses (Table 4).

Using either their total problem scores or their specific subscales, both SDQ versions (self-report or parent-rated) were similarly effective in predicting the respective clinical categories as the YSR or CBCL. Thus, the total difficulties score of the SDQ self-report was just as able to distinguish between patients with vs. without any psychiatric diagnosis (AUC = 0.835) as was the YSR total problems score (AUC = 0.810), as shown in the left column of Table 4. Moreover, it was seen that the subscales

Table 3 Correlations between SDQ scales from different informants (11- to 17-year-old Göttingen in- and outpatients with available parent, teacher, and self-report SDQs, N = 124)

SDQ scale	self x parent ¹ report	male ¹ (N = 83)	female ¹ (N = 41)	p ²	self x teacher ¹ report	male ¹ (N = 83)	female ¹ (N = 41)	p ²
Total Difficulties	0.39	0.29	0.61	*	0.27	0.21	0.42	ns
Emotional Symptoms	0.30	0.31	0.27	ns	0.29	0.20	0.30	ns
Conduct Problems	0.51	0.42	0.66	*	0.50	0.48	0.59	ns
Hyperactivity/Inattention	0.44	0.33	0.62	*	0.38	0.31	0.47	ns
Peer Problems	0.57	0.49	0.70	*	0.46	0.48	0.39	ns
Prosocial Behaviour	0.46	0.37	0.60	ns	0.32	0.21	0.41	ns

¹ all correlation coefficients significant at $p \leq 0.05$ (one-tailed Spearman rank correlations)

² * $p \leq 0.05$; ns not significant (one-tailed z-test comparing correlation coefficients obtained for male vs. female subgroups)

Table 4 Prediction of clinical diagnosis by parent- and self-rated SDQ or CBCL/YSR scales (AUC: Area under the curve = cut-off-independent area under the ROC curve; N = 214 Göttingen in- and outpatients with parent reports and self-ratings, including N = 20 without any diagnosis on axis I)

Predicted clinical target group	Considered questionnaire scale	Self-report		Parent		p ²
		AUC	p ¹	AUC	p ¹	
Any diagnosis on axis I yes: N = 194 vs. no: N = 20	SDQ – Total Difficulties Score	0.835		0.738		*
	YSR/CBCL – Total Problems Score	0.810	ns	0.746	ns	ns
Emotional disorders yes: N = 100 vs. no: N = 114	SDQ – Emotional Symptoms	0.690		0.712		ns
	YSR/CBCL – Internalising Problems	0.698	ns	0.728	ns	ns
Oppositional/conduct disorders yes: N = 69 vs. no: N = 145	SDQ – Conduct Problems	0.773		0.824		ns
	YSR/CBCL – Externalising Problems	0.742	ns	0.832	ns	**
Hyperactivity/attention-deficit disorders yes: N = 43 vs. no: N = 171	SDQ – Hyperactivity/Inattention	0.684		0.759	*	*
	YSR/CBCL – Attention Problems	0.623	ns	0.707		*

** $p \leq 0.01$; * $p \leq 0.10$; ns not significant

¹ two-tailed z-test comparing AUC differences between instruments (parent SDQ vs. CBCL; self-rated SDQ vs. YSR)

² two-tailed z-test comparing AUC differences between informants (parent- vs. self-rated SDQ; CBCL vs. YSR)

of the self-rated SDQ could effectively detect clinical subgroups of patients, particularly those with oppositional/conduct disorders. Again, there were only very small differences between the AUC measures obtained with the two self-report instruments, demonstrating a similar diagnostic efficiency.

Comparison between the predictive values of self-reports and parent ratings revealed a slight superiority of the self-rated SDQ, since its total difficulties score was marginally better at predicting the presence of any psychiatric disorders than the corresponding overall total score obtained from parent reports. Parent ratings, on the other hand, seemed to be slightly better predictors of more specifically defined diagnostic subgroups than self-reports, particularly concerning patients with clinically diagnosed hyperactivity and inattention disorders. This tendency was observed for both SDQ and YSR/CBCL instruments (right column of Table 4).

■ Predictive value of different combinations of SDQ informants

A different methodological approach, used to evaluate the relative contribution of individual SDQ informants to the overall prediction of clinical diagnoses, involved comparisons of adjusted R^2 values derived from regression analyses. Each adjusted R^2 shown in Table 5 reflects the proportion of variance of the dependent variable (i. e. subsamples with or without the indicated clinical diagnosis) that can be predicted using SDQ information from one or more sources.

As shown in Table 5, information from parents and teachers as well as from the children and adolescents themselves were all helpful in predicting a patient's clinical status, although the overall accuracy of predicting presence vs. absence of any or a more specific clinical diagnosis was rather low within this clinical sample of older children and adolescents. When considering only

ratings from one single source, SDQ self-reports proved to be as good as the parent-rated version at predicting the overall presence of a psychiatric diagnosis. However, using only parent or teacher information to detect specific diagnostic subgroups, each of these sources yielded better results than self-ratings alone – the exception being the emotional disorder subcategory, which was best predicted by the parent-rated emotional symptoms subscale and less well detected by teachers.

When, in addition to parent reports, SDQ self-ratings were added as predictor, the proportion of explained variance increased significantly (e. g. for conduct/oppositional disorders: parent-only = 24 %, parent + self-report = 28 %). Similar results were found when adding the self-report to teacher ratings (emotional symptoms: teacher-only = 7 %, teacher + self-report = 12 %). However, using information from parent and teacher ratings as predictors (i. e. combining data from both adult informants), the percentage of explained variance was generally higher than combining self-reported information with either parent or teacher ratings (e. g. for conduct/oppositional disorders: parent + self-report = 28 %, teacher + self-report = 32 %, and parent + teacher = 35 %). While the combination of all three sources (i. e. using self-ratings in addition to information from both teacher- and parent-rated SDQs) yielded an additional slight increase in adjusted R^2 , none of these observed increments reached statistical significance.

Thus, it was demonstrated that the ability to predict the indicated diagnostic subgroups was significantly improved when the self-report version was added to either parent or teacher ratings. If information from both adult sources (i. e. parents and teachers) was already available, the additional inclusion of self-ratings did not provide any improvement in predictive power.

Table 5 Prediction of diagnostic categories using different combinations of SDQ informants (adjusted R^2 values representing % of predicted variance of the respective target category; core sample of 11- to 17-year-old Göttingen in- and outpatients with complete SDQ data from all three informants; N = 124)

Diagnostical categories/scales (N: yes vs. no)	self only	parent only	parent + self	p	teacher only	teacher + self	p	parent + teacher	parent + teacher + self	p
Any diagnosis on axis I (N = 112 vs. 12) predicted by: Total difficulties score	0.10	0.09	0.12	*	0.13	0.17	*	0.15	0.17	ns
Emotional disorders (N = 55 vs. 69) predicted by: Emotional symptoms scale	0.09	0.17	0.20	*	0.07	0.12	*	0.20	0.21	ns
Conduct/oppositional disorders (N = 46 vs. 78) predicted by: Conduct problems scale	0.18	0.24	0.28	*	0.30	0.32	*	0.35	0.36	ns
Hyperactivity/attention-deficit disorders (N = 28 vs. 96) predicted by: Hyperactivity/Inattention scale	0.09	0.12	0.14	*	0.12	0.15	*	0.16	0.17	ns

* $p \leq 0.05$; ns not significant (two-tailed z-test of the effect of adding self-report data to parent and/or teacher ratings)

Discussion

The self-report version of the German SDQ was evaluated in a clinical sample of 214 in- and outpatients, aiming to determine the psychometric properties of this self-rated questionnaire and to investigate the diagnostic utility of adding self-reported information to external ratings by parents and teachers. Evaluation of the homogeneity of the total score and individual subscales demonstrated satisfactory reliability. In comparison with the corresponding scales of the parent version, SDQ self-ratings showed higher reliability for the emotional symptoms subscale, while, in contrast, other scales (conduct problems, hyperactivity-inattention, and peer problems) were more homogeneous in parent ratings.

Among the investigated patients, age effects were observed for some subscales of the SDQ self-report, older adolescents reporting more emotional symptoms and more prosocial behaviour than younger children. In line with the results of a previous clinical study [13], patients in the present clinical sample rated themselves as having fewer behavioural problems than according to their parents' reports.

Cross-informant correlations of self-ratings were found to be higher with parent than with teacher reports, replicating findings from other studies (e. g. [3]). Compared to boys, self-reports of girls showed considerably stronger agreement with their parents concerning peer relations, hyperactivity, inattention, and conduct problems.

The self-rated SDQ also turned out to possess at least comparable discriminative validity as the well-established but much longer YSR, both questionnaires being equally able to differentiate between patients with and without child psychiatric diagnoses. Further comparisons between the SDQ self-report and the corresponding YSR scales failed to find substantial differences in effectively detecting specific diagnostic subgroups.

SDQ self-reports were slightly better than the parent SDQ at identifying subgroups with and without any clinical diagnosis, while the parent version allowed a more accurate prediction of hyperactivity and attention-deficit disorders. Similar results have previously been reported by Klasen and coworkers [19].

The potential benefits of additionally considering self-report questionnaires to improve prediction of clinical status were investigated using a multiple regression approach. Here, SDQ scores as reported by either one, two, or all three informants served to predict presence of any child psychiatric diagnosis at all and, at a more specific level, broader-band diagnostic subcategories. When self-report information was added to either parent *or* teacher ratings, prediction of the clinical status improved. If, however, information from both other sources (i. e. parents and teachers) was already available, the additional inclusion of self-report data was unable to provide any further predictive power. It should be noted that, within this clinical sample, overall prediction of presence vs. absence of psychiatric diagnoses was rather poor, presumably reflecting an abundance of behavioural problems even in the fewer cases who did not fully meet the criteria for any or a specific clinical diagnosis.

In conclusion, the present results demonstrate that SDQ self-ratings from children and adolescents represent a useful contribution to the diagnostic process. Comparing the described findings obtained for the German SDQ self-report with those from other countries, our results on scale means, internal consistency, cross-informant correlations, and discriminative validity agree well with reports from England [13], the Netherlands [24], and Finland [20], thus further establishing the self-rated version of this easily handled and economical instrument as a powerful and reliable tool for a multitude of clinical and research applications.

References

1. Achenbach TM (1991) Manual for the Child Behavior Checklist/4–18 and 1991 Profile. University of Vermont, Department of Psychiatry, Burlington, VT
2. Achenbach TM (1991) Manual for the Youth Self-Report and 1991 Profile. University of Vermont, Department of Psychiatry, Burlington, VT
3. 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
4. Becker A, Woerner W, Hasselhorn M, Banaschewski T, Rothenberger A (2004) Validation of the parent and teacher SDQ in a clinical sample. *Eur Child Adolesc Psychiatry* 13(Supplement 2): 11–16
5. Bettge S, Ravens-Sieberer U, Wietzker A, Holling H (2002) Ein Methodenvergleich der Child Behavior Checklist und des Strengths and Difficulties Questionnaire. *Gesundheitswesen* 64: 119–124
6. Broberg AG, Ekeröth K, Gustafsson PA, Hansson K, Hägglöf B, Ivarsson T, Larsson B (2001) Self-reported competencies and problems among Swedish adolescents: a normative study of the YSR. *Eur Child Adolesc Psychiatry* 10: 186–193
7. Costello AJ, Edelbrock CS, Dulcan MK, Kalas R, Klaric SH (1984) Report on the NIMH Diagnostic Interview Schedule for Children (DISC). National Institute of Mental Health, Washington, D.C.
8. Döpfner M, Berner W, Lehmkuhl G (1995) Reliabilität und faktorielle Validität des Youth Self-Reports der Child Behavior Checklist bei einer klinischen Stichprobe. *Diagnostica* 41:221–244

9. Fergusson DM, Horwood LJ, Lynskey MT (1993) Prevalence and comorbidity of DSM-III-R diagnoses in a birth cohort of 15 year olds. *J Am Acad Child Adolesc Psychiatry* 32:1127–1134
10. Fombonne E (1991) The use of questionnaires in child psychiatry research: measuring their performance and choosing an optimal cut-off. *J Child Psychol Psychiatry* 32:677–693
11. Goodman R (1997) The Strengths and Difficulties Questionnaire: a research note. *J Child Psychol Psychiatry* 38: 581–586
12. 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
13. Goodman R, Meltzer H, Bailey V (1998) The Strengths and Difficulties Questionnaire: a pilot study on the validity of the self-report version. *Eur Child Adolesc Psychiatry* 7:125–130
14. Görtz A, Döpfner M, Nowak A, Bonus B, Lehmkuhl G (2002) Are self-reports of adolescents helpful in the assessment of Attention Deficit/Hyperactivity Disorders? An analysis with the assessment system DISYPS. *Kindheit und Entwicklung* 11:82–89
15. Hanley JA, McNeil BJ (1983) A method of comparing the areas under receiver operating characteristic curves derived from the same cases. *Radiology* 148: 839–843
16. Harrington R, Rutter M, Fombonne E (1996) Developmental pathways in depression: multiple meanings, antecedents, and endpoints. *Devel Psychopathol* 8:601–616
17. Hodges K (1993) Structured interviews for assessing children. *J Child Psychol Psychiatry* 34:49–68
18. Kaufman J, Birmaher B, Brent D, Rao U, Flynn C, Moreci P, Williamson D, Ryan N (1997) Schedule for Affective Disorders and Schizophrenia for School-Age Children – Present and Lifetime Version (K-SADS-PL): initial reliability and validity data. *J Am Acad Child Adolesc Psychiatry* 36:980–988
19. Klasen H, Woerner W, Wolke D, Meyer R, Overmeyer S, Kaschnitz W, Rothenberger A, Goodman R (2000) Comparing the German versions of the Strengths and Difficulties Questionnaire (SDQ-Deu) and the Child Behavior Checklist. *Eur Child Adolesc Psychiatry* 9:271–276
20. Koskelainen M, Sourander A, Kaljonen A (2000) The Strengths and Difficulties Questionnaire among Finnish school-aged children and adolescents. *Eur Child Adolesc Psychiatry* 9:277–284
21. Koskelainen M, Sourander A, Vauras M (2001) Self-reported strengths and difficulties in a community sample of Finnish adolescents. *Eur Child Adolesc Psychiatry* 10:180–185
22. Kvernmo S, Heyerdahl S (1998) Influence of ethnic factors on behavior problems in indigenous Sami and majority Norwegian adolescents. *J Am Acad Child Adolesc Psychiatry* 37:743–751
23. Morgan CJ, Cauce AM (1999) Predicting DSM-III-R disorders from the Youth Self-Report: analysis of data from a field study. *J Am Acad Child Adolesc Psychiatry* 38:1237–1245
24. Muris P, Meesters C, van den Berg F (2003) The Strengths and Difficulties Questionnaire (SDQ) – Further evidence for its reliability and validity in a community sample of Dutch children and adolescents. *Eur Child Adolesc Psychiatry* 12:1–8
25. Roth M (2000) Evaluation of the Youth Self-Report in a nonclinical sample. *Zeitschrift für Differentielle und Diagnostische Psychologie* 21:105–110
26. Saunders SM, Resnick MD, Hoberman HM, Blum RW (1994) Formal help-seeking behavior of adolescents identifying themselves as having mental health problems. *J Am Acad Child Adolesc Psychiatry* 33:718–728
27. Sawyer MG, Baghurst P, Mathias J (1992) Differences between informants' reports describing emotional and behavioural problems in community and clinic-referred children: a research note. *J Child Psychol Psychiatry* 33: 441–449
28. Steinhausen H-C, Winkler-Metzke C (1998) Youth Self-Report of behavioral and emotional problems in a Swiss epidemiological study. *J Youth Adolesc* 27: 429–441
29. Unnewehr S, Schneider S, Margraf J (1995) Diagnostisches Interview psychischer Störungen im Kindes- und Jugendalter. Springer, Berlin
30. Verhulst FC, van der Ende J (1992) Agreement between parents' reports and adolescents' self-reports of problem behavior. *J Child Psychol Psychiatry* 33:1011–1023
31. Woerner W, Becker A, Friedrich C, Klasen H, Goodman R, Rothenberger A (2002) Normierung und Evaluation der deutschen Elternversion des Strengths and Difficulties Questionnaire (SDQ): Ergebnisse einer repräsentativen Felderhebung. *Zeitschrift für Kinder- und Jugendpsychiatrie und Psychotherapie* 30:105–112
32. Zwaanswijk M, van der Ende J, Verhaak PF, Bensing JM, Verhulst FC (2003) Factors associated with adolescent mental health service need and utilization. *J Am Acad Child Adolesc Psychiatry* 42: 692–700