

Validity of the social responsiveness scale to differentiate between autism spectrum disorders and disruptive behaviour disorders

Hannah Cholemker · Janina Kitzerow ·
Sonja Rohrmann · Christine M. Freitag

Received: 4 December 2012 / Accepted: 15 May 2013 / Published online: 30 May 2013
© Springer-Verlag Berlin Heidelberg 2013

Abstract Autism spectrum disorder (ASD) as well as oppositional defiant (ODD) and conduct disorder (CD) is characterised by difficulties in social interaction with peers. The Social Responsiveness Scale (SRS) measures reciprocal social behaviour in children and adolescents and was originally developed as a quantitative measure of autistic traits. In the present study, we compare parent-rated SRS scores in children with ODD, CD, and ASD and examine the diagnostic validity of the SRS alone and in combination with additional questionnaires to differentiate between groups. We hypothesize that the SRS better differentiates ASD and typically developing controls (TD) than ASD and the disruptive behaviour disorders ODD and CD. The sample consists of three clinical groups: ASD without comorbid intellectual delay ($N = 55$), ODD/CD ($N = 55$), and TD ($N = 55$), between 6 and 18 years. The groups were matched by age, sex, and IQ. SRS scores were compared for the three groups. Sensitivity and specificity of the SRS total and sub-scores were examined by receiver operating characteristics (ROC) analyses. Logistic regression analyses were calculated for estimating the rate of correctly specified individuals. The SRS differentiated excellently between ASD and TD (ROC-AUC = 1.00), but sensitivity and specificity were considerably lower when ASD was compared with ODD/CD (ROC-AUC = 0.82). A combination of three parent-rated questionnaires resulted

in an improved validity to differentiate ASD and ODD/CD. For clinical screening purposes in children suspicious of ASD and/or ODD/CD, the SRS should be used in combination with additional disorder-specific questionnaires to improve the rate of correct classification of both disorders.

Keywords Psychometric assessment · Differential diagnosis · Child psychiatric disorder

Introduction

Autism spectrum disorders (ASD) are characterised by severe and pervasive impairments in three domains: reciprocal social interaction and communication skills and the presence of stereotyped behaviour, interests, and activities [4, 42]. The three subtypes (autistic disorder, Asperger's disorder and pervasive developmental disorder—not otherwise specified) differ with regard to symptom pattern, severity, and associated cognitive and language abilities. The disorders are currently conceptualised to lie on a continuum of autism-specific traits. This will be taken on in the 5th edition of the Diagnostic and Statistical Manual of Mental Disorders (<http://www.dsm5.org>), which will contain only one diagnosis of ASD with varying degrees of severity.

The current gold standard to diagnose ASD is a combination of the Autism Diagnostic Interview-Revised (ADI-R) and the Autism Diagnostic Observation Schedule (ADOS) combined with clinical judgement [32]. However, both ADI-R and ADOS are very time-consuming and an extensive training is necessary, which limits their feasibility in clinical settings [17]. In contrast, parent-rated questionnaires are quickly accomplishable, objective, economic and easy to apply. Thus, several questionnaires have been developed which aim at screening children and

H. Cholemker (✉) · J. Kitzerow · C. M. Freitag
Department of Child and Adolescent Psychiatry, Psychosomatics
and Psychotherapy, JW Goethe University Hospital,
Deutschordenstraße 50, 60528 Frankfurt am Main, Germany
e-mail: Hannah.Cholemker@kgu.de

S. Rohrmann
Department of Psychology, JW Goethe University,
Frankfurt am Main, Germany

adolescents for ASD, and then confirm the diagnosis in positively screened individuals by ADI-R and ADOS [39].

The Social Responsiveness Scale (SRS) is a recently developed ASD questionnaire which is increasingly being used as a clinical screening instrument. [14]. Whereas other ASD screening questionnaires like the Social Communication Questionnaire (SCQ) [34] have been derived from the ADI-R with the categorical concept of DSM-IV TR or ICD-10 diagnostic criteria, the SRS was developed as a measure of quantitative autistic traits in children and adolescents. It showed a single-factor solution with a continuously distributed score and was originally validated and used in population-based samples [14–16]. The SRS is not only focussed on assessing reciprocal social behaviour, but also includes core symptoms of ASD as communication deficits and stereotyped behaviour as well as symptoms which are not exclusively related to ASD [7, 23]. The majority of the items were also deduced from DSM-IV TR ASD symptoms, but some items were selected additionally as less specific but frequently observed symptoms in children and adolescents with ASD [7].

For a good screening tool it is essential that it not only picks up individuals with ASD correctly (sensitivity) but also does not overestimate ASD in children with other psychiatric disorders or in typically developing individuals (specificity) [18]. To assess and compare the discriminant power in distinguishing ASD from other psychiatric disorders, receiver operating characteristics (ROC) with the respective area under the curve (AUC), sensitivity and specificity are calculated as a commonly used method of choice [7, 10, 13, 18]. Accuracy of the test is measured by the AUC. An AUC of 1 (100 %) indicates an optimal test with maximum sensitivity and specificity, whereas an area of 0.5 (50 %) reflects a chance finding. In addition, the AUC is used to establish an optimal cut-off score with maximum sensitivity and specificity.

When used as a screening tool for ASD, the parent-rated SRS differentiated well between ASD and TD (AUC = 0.98) [12]. However, when the screening properties of the SRS were compared between children with ASD and children with other unspecific psychiatric disorders, sensitivity and specificity were considerably lower [10]. Constantino and Gruber [14] reported a sensitivity of 0.70 and a specificity of 0.90 (SRS total cut-off score of 85) for children with ASD aged 4 to 18 years old compared with another clinical sample (AUC = 0.85). In a German study of 480 participants (aged 4–18 years) the ROC showed an AUC = 0.98 for an SRS total cut-off score of 85 when ASD and TD were compared, whereas AUC declined to 0.81 when ASD had to be differentiated from mixed child psychiatric disorders [12]. In a British sample of 119 children aged 9–13 years, Charman et al. [13] reported a somewhat higher sensitivity (0.78) but a reduced specificity (0.67) with a cut-off score of 75 in children with special educational needs with and

without ASD (AUC = 0.77). Efficacy was especially low for intellectually disabled children and for children with additional behaviour problems (AUC_{both} = 0.67).

Assessment of convergent validity in the German sample resulted in positive, but rather low, correlations with the SCQ total, ADI-R and ADOS-derived scores. The highest correlations were shown with the SCQ ($r = 0.58$) [10]; lower correlations were demonstrated for ADI-R ($r = 0.46$ for social interaction, $r = 0.40$ for communication and $r = 0.38$ for stereotyped behaviour [10]). For ADOS a correlation of $r = 0.35$ for the social and communication score was reported [10]. Concerning concurrent validity, the SRS showed moderate to strong correlations with the Child Behaviour Checklist (CBCL) sub-scales “Thought Problems”, “Social Problems”, “Attention Problems”, “Aggressive Behaviour” and the CBCL overall scale “total score” [7, 16].

Although the SRS was developed to assess autistic traits, many items describe symptoms which are not exclusively related to ASD, e.g. “anxious in social interaction”, “is suspicious” or “poor concept of cause and effect” [22]. Several studies reported a rather unspecific positive correlation of non-ASD behaviour problems and SRS scores [13, 24]. Hus et al. [24] even suggested that increased SRS scores may better be interpreted as indicator of a general level of impairment than as severity score of ASD specific symptoms. Especially children with anxiety disorders, oppositional defiant and conduct disorders are particularly impressive examples of mental health problems that are strongly associated with difficulties in social interaction and reactivity. Previous studies analysed sensitivity and specificity in mixed clinical samples in children with ADHD or anxiety disorder, but not in children with ODD or CD. Interestingly, however, children with ODD or CD reached the second highest SRS scores below children with ASD in one study [7, 10]. ODD and CD symptoms are very closely correlated throughout development and show a shared underlying genetic and environmental liability [21]. In addition, impairments in social interaction, difficulties in finding friends, misperceiving intentions of others, little concern for feelings of others, aggressive and oppositional behaviour, reduced frustration tolerance as well as irritability [39] or lack of empathy [37] are characteristics not only of children with both, ODD and CD, but also of ASD. Comparable difficulties of ODD/CD and ASD in reciprocal social interaction have also been described [2]. One study reported that a large number of children later diagnosed with ASD previously had received a diagnosis of ADHD (21.4 %) or ODD/CD (12 %) [29], likely due to overlapping difficulties in social interaction. On the other hand, children and adolescents diagnosed with ASD often also show a high rate of concomitant behaviour problems, especially ODD and ADHD symptoms [25–27], which also complicates the correct differential diagnosis of ODD, CD,

and ASD, even more so in children without intellectual disability. With focus on the overlapping symptomatology the aim of this study was to evaluate the screening accuracy of the SRS in differentiating ASD and ODD/CD. As no previous study specifically has compared these groups, we here aimed first at assessing the diagnostic validity of the SRS to differentiate ASD from the disruptive behaviour disorders ODD and CD, which are also characterised by difficulties in reciprocal social interaction. In order to support time-efficient clinical diagnostic procedures [18], we tested if the rate of correct classifications by the SRS could be improved by additional standardized parent questionnaires. Additionally, we aimed at replicating the above-mentioned findings on convergent (SCQ, ADI-R) and concurrent (CBCL) validity of the SRS.

We compared parent-rated SRS scores in 6- to 18-year-old children and adolescents with ODD/CD, ASD without comorbid intellectual delay, and well-matched TD, and studied the diagnostic validity of the SRS to differentiate between these groups. For this purpose we used ROC-analyses to elicit if it was possible to distinguish the groups with an optimal sensitivity and specificity by SRS.

With regard to the SRS total and the five different SRS subscales we expected that children and adolescents with ASD will score higher than children and adolescents with ODD/CD or TD and that children and adolescents with ODD/CD will score higher than TD. We hypothesized that the SRS will better differentiate ASD and TD than ASD and ODD/CD.

We also expected an increased correct classification rate by adding the SCQ- and an ODD/CD-specific German screening questionnaire based on DSM-IV TR and ICD-10 ODD and CD symptoms (FBB-SSV).

Methods

Participants

The entire sample included 165 children and adolescents aged 6–18 years, 55 individuals with ASD, 55 individuals

with ODD or CD diagnosed according to DSM IV-TR, as well as 55 TD. To ensure a better comparability, the groups were matched for IQ, age and gender (Table 1). Diagnoses in the clinical groups were established by independent clinicians (psychologists, psychiatrists) approximately 1–6 months prior to the questionnaire-based study. Due to the recruitment of all clinical participants as outpatients, a moderate and fairly equally distributed degree of severity can be assumed for both, ASD and ODD/CD.

The ASD sample included 40 individuals with Autism and 15 with Asperger's Syndrome, 47 (85 %) males, and 8 (15 %) females with an average IQ of 100.6 (SD 15.2), and a mean age of 12.5 years (SD 2.7). Only those participants fulfilling diagnostic and research criteria of ASD by a combination of ADI-R, and ADOS were included. The disruptive behaviour disorder group consisted of 37 (67 %) individuals with CD and 18 (33 %) with ODD, 45 (82 %) males and 10 (18 %) females; the average IQ was 98.5 (SD 13.5), the mean age 12.4 years (SD 3.1), diagnosed by clinician experts, and confirmed by K-Dips. The typically developing group included 45 (82 %) male and 10 (18 %) female participants, with an average IQ of 103.4 (SD 14.5) and a mean age of 11.9 years (SD 2.9). This group had no psychiatric symptoms according to the CBCL [1].

The study was approved by the ethical committee of the medical faculty, Goethe University, Frankfurt am Main, Germany. Informed consent was obtained from the parents and children. Clinical groups were studied in the year 2011 after completion of the diagnostic process before any kind of treatment. The Department of Child and Adolescent Psychiatry, Psychosomatics, and Psychotherapy, Goethe University, Frankfurt am Main has an outpatient clinic which serves all local children and adolescents referred for diagnosis and treatment of psychiatric disorders. Among others, research focus concentrates on ASD and ODD/CD. The typically developing participants were recruited from local schools and advertisements. All participants received a moderate fee for participation.

Table 1 Sample description: gender, age, and IQ

	ASD (<i>N</i> = 55)	CD/ODD (<i>N</i> = 55)	TD (<i>N</i> = 55)	Statistical group differences		Post hoc
				<i>F</i> / χ^2 value (<i>df</i>)	<i>p</i> value	
Gender						
Female/male	8/47	10/45	10/45	0.344 (2)	0.842	–
Age						
<i>M</i> (SD)	12.5 (2.7)	12.4 (3.1)	11.9 (2.9)	0.831 (2/162)	0.437	–
IQ						
<i>M</i> (SD)	100.6 (15.2)	98.8 (13.5)	103.4 (14.5)	0.944 (2/126)	0.392	–

ASD autism spectrum disorder, CD conduct disorder, ODD oppositional defiant disorder, TD typically developing, IQ intelligence quotient, *N* sample size, *M* mean, *SD* standard deviation

Measures

The German version of the ADI-R [11, 35] and the ADOS [28, 33] was performed with all ASD individuals to confirm the clinical diagnosis by clinical experts (psychologists, psychiatrists) who were trained to research standards. Both are excellently validated diagnostic tools, based on ICD-10/DSM-IV-TR criteria. The ADI-R is a detailed interview with the primary caretaker on lifetime ASD symptoms. The ADOS is a direct observation schedule with four different modules for children, adolescents, and adults with varying developmental age and language abilities. In this study, modules 2, 3, and 4 were used. ADI-R and ADOS provide empirically derived diagnostic algorithms for each of the three subdomains of autism: social interaction, communication, and stereotyped behaviours.

Primary and comorbid psychiatric diagnoses in children and adolescents with ODD/CD and ASD were obtained from *The Diagnostic Interview for Children and Adolescents*, parent version (Kinder-DIPS) for almost all individuals of the clinical sample ($n = 103$) [36]. $N = 7$ parents declined participation in the Kinder-DIPS. The Kinder-DIPS is a structured interview designed to assess common mental disorders in children and adolescents according to ICD-10 and DSM-IV-TR criteria. Symptom frequency and/or severity is assessed on a four-point Likert scale varying from 0 (never) to 4 (very often). The Kinder-DIPS is widely used in German-speaking populations and has shown good retest and inter-rater reliability [3].

The SRS [14] is a 65-item rating scale on social and autistic behaviour over the previous 6 months for 4- to 18-year-olds. It is a parent/teacher questionnaire and can be completed within 15–20 min. Each item is scored from 0 (never true) to 3 (almost always true), generating a total score in the range from 0 to 195. Total raw scores can be transformed into T scores. The manual recommends the use of SRS raw scores in research for comparability to previous studies of the SRS [7, 24]. Scores can also be obtained for five sub-scales: social awareness, social cognition, social communication, social motivation, and autistic mannerisms. The German adaption [7] was used in this study. Consistent with the English original, this version has demonstrated good to excellent psychometric properties in the German standardization study [10].

In addition, the SCQ [6, 34] was obtained. It is a 40-item parent-report screening questionnaire for autism, based on the ADI-R, with good psychometric properties [8].

Psychiatric symptoms in different domains were assessed with the German version of the CBCL 4-18 [1, 20], an internationally validated and widely used parent-report form with 113 items, computing a total score, overall scores for internalizing and externalizing problems, and syndrome scales for various behavioural and emotional

problems (withdrawn, somatic complaints, anxious/depressed, social problems, thought problems, attention problems, delinquent behaviour and aggressive behaviour). Responses are recorded on a Likert Scale from 0 (not true), 1 (sometimes true) up to 2 (often true). We calculated correlations between the SRS and the CBCL total scores, the two-second order, and the eight-syndrome scales.

Three parent questionnaires from the *Diagnostic System for Mental Disorders in Children and Adolescents* (DISYPS-II) [19], i.e. German diagnostic symptom checklists according to DSM-IV TR and ICD-10, were used to additionally quantify ODD/CD symptoms (25 items), anxiety and obsessive–compulsive behaviours (33 items) and ADHD symptoms (20 items). The parent questionnaire for ODD/CD (FBB-SSV) includes 9 items of ODD and 16 items of CD symptoms. Anxiety and obsessive–compulsive symptoms (FBB-ANZ) are obtained for four anxiety disorders: separation anxiety (10 items), generalized anxiety (7 items), social phobia (7 items), and specific phobia (7 items), and two additional items for OCD symptoms. ADHD symptoms (FBB-ADHS) are measured on three scales: attention problems (9 items), hyperactivity (7 items), and impulsivity (4 items). Symptoms are rated on a Likert scale from 0 (no problems) to 3 (most severe problem). Validation studies of the DISYPS-II have demonstrated good reliability ($\alpha = 0.71$ to 0.94) and appropriate validity for the parent rating questionnaires [19]. In the present study, the raw total score of each of the three questionnaires (FBB-SSV, FBB-ANZ, FBB-ADHS) was used to compare groups and explore their diagnostic validity to separate the groups.

All questionnaire data (SRS, SCQ, CBCL, FBB-SSV, FBB-ADHS, FBB-ANZ) were obtained from the parents after the respective diagnoses (ASD or ODD/CD) had been established. Questionnaire data were analysed by an independent researcher not involved in the diagnostic process using in-house computer-based calculation algorithms based on the respective manuals.

IQ measures were assessed by the age-appropriate German version of the Wechsler Intelligence Scales for children and adolescents (Hamburg-Wechsler-Intelligence Test for children, HAWIK-IV) [31] and adults (Wechsler Intelligence Test for adults, WIE) [5] or the current version of the revised Culture Fair Intelligence Test (CFT 20-R) [41].

Data analysis

Descriptive measures were compared by Pearson χ^2 -test, parametric or non-parametric ANOVA, as appropriate. Frequency distribution of gender was assessed by χ^2 -test, group differences in age and IQ were tested by ANOVA followed by Scheffé tests after verifying normal

distribution with Kolmogorov-Smirnov-Test and variance homogeneity with Levené-Test. The rates of comorbid diagnoses were compared by χ^2 -test. Correlation analysis was done by Pearson correlation, as the respective data (IQ, age, questionnaire derived measures) met the normal distribution assumption. In case of normal distribution violations (e.g. SRS sub-scales), Spearman correlations were calculated.

Group differences of the SRS total score were compared by ANOVA and subsequent Scheffé-Tests to test for pairwise differences. Significance level was set at $\alpha = 0.05$ (uncorrected). With a power of $1 - \beta \geq 0.8$, a medium effect of $\delta = 0.25$ can be observed by ANOVA in a three-group sample of 165 (3×55) individuals. Due to normal distribution violations, group differences of the SRS sub-scales were compared by the non-parametric Kruskal-Wallis test followed by Mann-Whitney U tests for pairwise comparisons.

Diagnostic validity was analysed by the receiver operating characteristics curve (ROC) comparing ASD vs. ODD/CD and ASD vs. TD for the SRS total score and sub-scales. ROC illustrates the performance of a binary classifier system. The AUC, sensitivities (true positives) and specificities (1-true negatives) for optimal SRS cut-offs, determined by the Youden-score, were calculated. Test accuracy is measured by AUC with an area of 1 representing perfect classification and an area of 0.5 showing a random result. Similar analyses (group differences, ROC-AUC calculation) were also performed with the CBCL, FBB-SSV, FBB-ANZ, FBB-ADHD and the SCQ, to assess their validity to differentiate among the three groups. For a better translation of the meaning of the ROC-AUC results into everyday diagnostics, the predictive value of a positive (PV+) and a negative test (PV-) was additionally calculated for the SRS scores (at the optimal cut-off) in the clinical groups [18]. Whereas PV+ represents the percentage of patients with a positive test result who actually have the condition, PV- gives the probability that the patient really does not have the condition, when the test is negative.

Logistic regression is generally used to predict the odds of being a case based on predictor variables. It can also be used to study the correct classification into two groups by the predictor variables. Thus, to compare if the SRS total score as predictor alone or in combination with the predictor variables SCQ and/or FBB-SSV resulted in the best correct classification of ASD and ODD/CD individuals, binary logistic regression analyses were done. At the respective cut-off dichotomised questionnaire data were included stepwise into the model: first the SRS, then the SCQ and FBB-SSV predicting the ASD (ASD = 1) versus the ODD/CD (ODD/CD = 0) group. Contribution of the predictors to the model was tested by the Chi-squared Wald

statistic. To compare two different models, we used the likelihood ratio test.

Convergent validity of the SRS was explored by Spearman correlation of the SRS total and sub-scores with ADI-R and the SCQ total score because of normal distribution violations. Concurrent validity of the SRS was similarly explored by Spearman correlation of the SRS total raw score with the CBCL sub- and total score.

Results

Descriptive measures: clinical information

Descriptive data on the three groups are shown in Tables 1 and 2. There were no significant differences in gender, age and intelligence between the three groups. The highest SRS and SCQ scores were found for ASD, followed by ODD/CD and TD. Comparing CBCL scales between groups, children with ODD/CD showed the highest scores in *externalizing, delinquent and aggressive behaviour*, whereas ASD children scored higher than ODD/CD in *internalizing behaviour, withdrawn, social problems and thought problems*. For the DISYPS-II scales, significant differences between clinical groups were found for *social phobia* with highest scores in ASD (FBB-ANZ), whereas *hyperactivity and total score* (FBB-ADHS) and both scales of the FBB-SSV were increased in ODD/CD (Table 2).

Both clinical groups ($N = 103$) showed a similar rate of comorbid diagnoses (Table 3) according to the Kinder-DIPS [35]. In the ASD group 54.9 % ($N = 28$) and in the ODD/CD group 55.8 % ($N = 29$) showed at least one comorbid diagnosis. Differences between groups were not significant ($\chi^2 = 0.533$, ns). Children with ASD most often also suffered from ADHD ($N = 15$; 29.4 %), tic disorders ($N = 6$; 11.8 %), and anxiety disorders ($N = 5$; 9.8 %); children with ODD/CD most often showed comorbid ADHD ($N = 21$; 40.4 %) followed by elimination disorders ($N = 4$; 7.7 %), sleeping disorders ($N = 4$, 7.7 %) and tic disorders ($N = 3$; 5.8 %).

SRS raw scores and moderating effects

Before comparing group differences, an influence of IQ, gender and age on SRS raw scores was excluded. IQ did not correlate strongly with the SRS total score across ($r = -0.14$, ns) and within groups (for ASD $r = -0.272$, ns; for ODD/CD $r = -0.094$, ns; for TD $r = 0.109$, ns). Comparable results were found for age across ($r = 0.15$, ns) and within groups (for ASD $r = 0.178$, ns; for ODD/CD $r = 0.097$, ns; for TD $r = 0.072$, ns). Females ($M = 63.3$, $SD = 42.4$) showed slightly higher SRS total raw scores than males ($M = 59.0$, $SD = 38.7$) across the three groups,

Table 2 Sample description: parent rated behavioural measures between all groups

	ASD (<i>N</i> = 55)	CD/ODD (<i>N</i> = 55)	TD (<i>N</i> = 55)	Statistical group differences		Post hoc
				<i>F</i> / χ^2 value (<i>df</i>)	<i>p</i> value	
SRS score						
M (SD)	97.0 (26.7)	62.7 (26.8)	19.4 (12.9)	155.97 (2/162)	0.000	ASD > ODD/CD > TD
Female	111.9 (25.7)	65.5 (26.1)	22.2 (15.4)			
Male	94.5 (26.3)	62.1 (27.2)	18.8 (12.5)			
	ASD (<i>N</i> = 50)	CD/ODD (<i>N</i> = 54)	TD (<i>N</i> = 51)	Statistical group differences		Post hoc
				<i>F</i> / χ^2 value (<i>df</i>)	<i>p</i> value	
SCQ score						
M (SD)	17.4 (7.8)	7.7 (5.1)	2.9 (2.9)	83.68 (2)	0.000	ASD > ODD/CD > TD
	ASD (<i>N</i> = 53)	CD/ODD (<i>N</i> = 47)	TD (<i>N</i> = 55)	Statistical group differences		Post hoc
				<i>F</i> / χ^2 value (<i>df</i>)	<i>p</i> value	
CBCL, <i>M</i> (SD)						
Total score	69.68 (6.87)	69.79 (8.15)	45.75 (8.17)	170.295 (2/154)	0.000	ASS/ODD/CD > TD
Internalizing	67.42 (7.23)	63.09 (8.83)	46.58 (6.80)	91.25 (2)	0.000	ASD>ODD/CD > TD
Externalizing	62.32 (8.65)	69.66 (7.81)	44.47 (7.84)	97.17 (2)	0.000	ODD/CD > ASD > TD
Withdrawn	69.68 (9.33)	63.36 (9.10)	51.35 (2.89)	85.51 (2)	0.000	ASD > ODD/CD > TD
Somatic complaints	60.53 (10.50)	59.13 (8.45)	53.13 (5.12)	21.51 (2)	0.000	ASS/ODD/CD > TD
Anxious/depressed	64.09 (7.66)	62.15 (7.19)	51.38 (2.87)	83.94 (2)	0.000	ASS/ODD/CD > TD
Social symptoms	72.68 (10.61)	65.85 (10.96)	51.91 (3.84)	91.5 (2)	0.000	ASD > ODD/CD > TD
Thought problems	67.53 (10.33)	60.70 (11.94)	50.96 (4.05)	63.95 (2)	0.000	ASD > ODD/CD > TD
Attention symptoms	70.00 (9.00)	69.62 (9.87)	52.16 (3.73)	89.14 (2)	0.000	ASS/ODD/CD > TD
Delinquent behaviour	59.85 (6.77)	66.26 (9.24)	51.33 (2.71)	78.85 (2)	0.000	ODD/CD > ASD > TD
Aggressive behaviour	63.75 (9.74)	71.79 (9.46)	51.31 (2.65)	90.42 (2)	0.000	ODD/CD > ASD > TD
	ASD (<i>N</i> = 38)	CD/ODD (<i>N</i> = 50)	TD (<i>N</i> = 32)	Statistical group differences		Post hoc
				<i>F</i> / χ^2 value (<i>df</i>)	<i>p</i> value	
FBB_ANZ, <i>M</i> (SD)						
Separation anxiety	0.50 (0.80)	0.76 (0.39)	0.16 (0.72)	8.26 (2)	0.016	ASS/ODD/CD > TD
Generalized anxiety	1.58 (1.80)	1.70 (1.61)	0.06 (0.25)	31.00 (2)	0.000	ASS/ODD/CD > TD
Social phobia	2.21 (1.80)	1.12 (1.47)	0.06 (0.35)	38.82 (2)	0.000	ASD>ODD/CD > TD
Specific phobia	0.66 (1.07)	0.52 (0.79)	0.03 (0.18)	13.21 (2)	0.001	ASS/ODD/CD > TD
Total	5.21 (4.11)	4.22 (3.75)	0.34 (0.94)	42.22 (2)	0.000	ASD/ODD/CD > TD
	ASD (<i>N</i> = 43)	CD/ODD (<i>N</i> = 55)	TD (<i>N</i> = 55)	Statistical group differences		Post hoc
				<i>F</i> / χ^2 value (<i>df</i>)	<i>p</i> value	
FBB_ADHD, <i>M</i> (SD)						
Attention problems	4.37 (2.95)	5.29 (2.71)	0.67 (1.50)	72.06 (2)	0.000	ODD/CD/ASD > TD
Hyperactivity	1.26 (1.81)	2.47 (2.45)	0.05 (0.30)	46.77 (2)	0.000	ODD/CD > ASD > TD
Impulsivity	1.28 (1.28)	1.80 (1.65)	0.11 (0.42)	46.09 (2)	0.000	ODD/CD/ASD > TD
Total	6.91 (4.78)	9.56 (5.66)	0.84 (1.70)	78.9 (2)	0.000	ODD/CD > ASD > TD

but differences were not significant ($U = 1823$, $p = 0.68$, $ES = 0.11$). In the ASD group, females showed a mean SRS total score of $M = 111.9$ ($SD = 25.7$), and males $M = 94.5$ ($SD = 26.3$; $t(53) = -1.74$; $p = 0.09$). Similar

results were found for ODD/CD (females $M = 65.5$, $SD = 26.1$, males $M = 62.1$ ($SD = 27.2$), $t(53) = -0.36$, $p = 0.72$), and TD (females $M = 22.2$, $SD = 15.4$, males $M = 18.8$, $SD = 12.5$, $t(53) = -0.74$, $p = 0.46$).

Table 2 continued

	ASD (<i>N</i> = 40)	CD/ODD (<i>N</i> = 50)	TD (<i>N</i> = 52)	Statistical group differences		Post hoc
				<i>F</i> / χ^2 value (<i>df</i>)	<i>p</i> value	
FBB_SSV, <i>M</i> (<i>SD</i>)						
Conduct disorder	0.18 (0.39)	1.58 (1.85)	0.06 (0.24)	87.15 (2)	0.000	ODD/CD > ASD > TD
Oppositional defiant disorder	2.08 (1.86)	4.52 (2.02)	0.19 (0.60)	49.41 (2)	0.000	ODD/CD > ASD; ODD/CD > TD

ASD autism spectrum disorder, CD conduct disorder, ODD oppositional defiant disorder, TD typically developed, *M* mean, *SD* standard deviation, SRS social responsiveness scale, SCQ social communication questionnaire, CBCL child behaviour checklist, FBB_ANZ external assessment for Anxiety and Obsessive–Compulsive behaviour, FBB_ADHD external assessment for Attention Deficit Hyperactivity Disorder, FBB_SSV external assessment for Conduct and Oppositional Defiant Disorder, > significant differences

Table 3 Psychiatric comorbid diagnoses in ASD and ODD/CD

	ASD <i>N</i> = 51, <i>N</i> (%)	ODD/CD <i>N</i> = 52, <i>N</i> (%)	Statistics	
			χ^2 value (<i>df</i>)	<i>p</i> value
Comorbidities (Kinder-Dips)				
Total ^a (<i>n</i> = 103),	28 (54.9)	29 (55.8)	0.533 (3)	0.912
ADHD	15 (29.4)	21 (40.4)	1.36 (1)	0.243
Anxiety disorders	5 (9.8)	2 (3.8)	1.44 (1)	0.230
Tic disorders	6 (11.8)	3 (5.8)	1.16 (1)	0.281
Enuresis/encopresis	2 (3.9)	4 (7.7)	0.67 (1)	0.414
Sleeping disorders	4 (7.8)	4 (7.7)	0.00 (1)	0.977
Others	1 (2.0)	4 (7.7)	1.83 (1)	0.176

^a Total means at least one diagnosis

ASD autism spectrum disorder, CD conduct disorder, ODD oppositional defiant disorder

SRS total raw scores strongly differed between groups ($F = 178.84, p < 0.001$). The highest scores were observed in ASD ($M = 97.0, SD = 26.7$), followed by ODD/CD ($M = 59.1, SD = 22.6$). The healthy control group scored lowest with $M = 19.4 (SD = 12.9)$. Similar findings were observed for the SRS sub-scales: ASD scored highest, followed by ODD/CD and healthy controls ($p_{all} < 0.001$). Within groups no differences were observed for ODD ($M = 63.2, SD = 26.0$) and CD ($M = 61.7, SD = 29.0$) and the two ASD diagnoses Autism ($M = 96.6, SD = 28.3$) and Asperger’s Syndrome ($M = 98.1, SD = 22.7$).

Discriminant validity of the SRS alone

An ROC space is defined by the area under the curve of a line plotting false-positive rates and true-positive rates against each other on related *x* and *y* axes, illustrating benefits (true-positive) and costs (false-positive). The best possible prediction consists of an ROC-AUC = 1, representing a perfect classification. An ROC curve near the

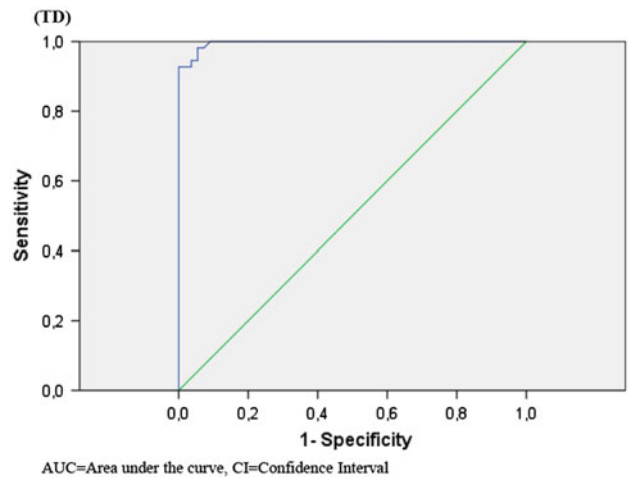


Fig. 1 Receiver Operating Characteristics (ROC) curve (AUC = 1.00, CI = 0.99–1.0) of the Social Responsiveness Scale (SRS) for Autism Spectrum Disorder (ASD) versus Typically Developing children (TD)

diagonal line depicts a random result. The SRS differentiated ASD and TD with an ROC-AUC = 1.0 (95 % CI 0.99–1.0) (Fig. 1), what means a nearly perfect classification. An SRS score of 43 showed the best sensitivity of 0.98 and specificity of 0.95, calculated by Youden-Score. Predicting ASD and ODD/CD by the SRS resulted in an ROC-AUC = 0.82 (95 % CI 0.74–0.90) (Fig. 2). An SRS total score of 80 showed the best sensitivity of 0.76 and specificity of 0.82. The predictive value (PV+) is the probability that a positive test result really identifies an individual with ASD. Here, the PV+ was 81 % for a SRS total score of 80. The predictive value of the negative test (PV–) showed a probability of 77 % that individuals really did not show the ASD condition when the test was negative.

Regarding the predictive value of the SRS sub-scales to differentiate between both disorders, the best performance was shown by the scales *autistic mannerisms* (ROC-AUC = 0.83, 95 % CI 0.75–0.91) and *social communication* (ROC-AUC = 0.83, 95 % CI 0.75–0.91). The other sub-scales *social awareness* (ROC-AUC = 0.71, 95 % CI

0.62–0.81), *social cognition* (ROC-AUC = 0.74, 95 % CI 0.65–0.83), and *social motivation* (ROC-AUC = 0.75, 95 % CI 0.65–0.84) showed a ROC-AUC < 0.80 (Fig. 3).

Discriminant validity of the SRS in combination with other questionnaires

Besides the SRS, the following parent rating scales showed differing mean scores between both clinical groups: CBCL (*internalizing, externalizing, withdrawn, social problems, thought problems, delinquent behaviour and aggressive behaviour*), the SCQ total score, the FBB-SSV (both sub-scales), FBB-ANZ (*social anxiety*),++ and FBB-ADHD

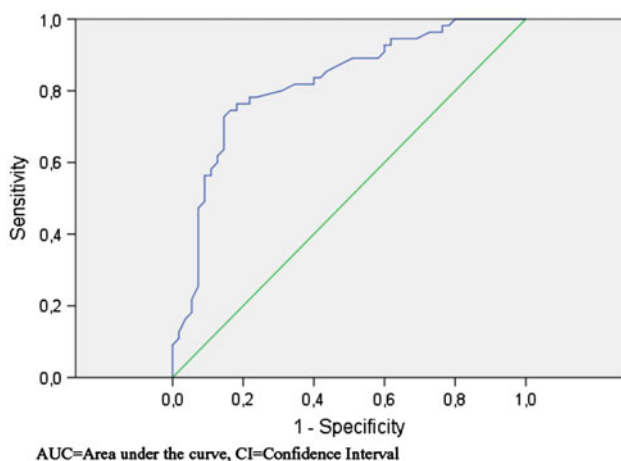
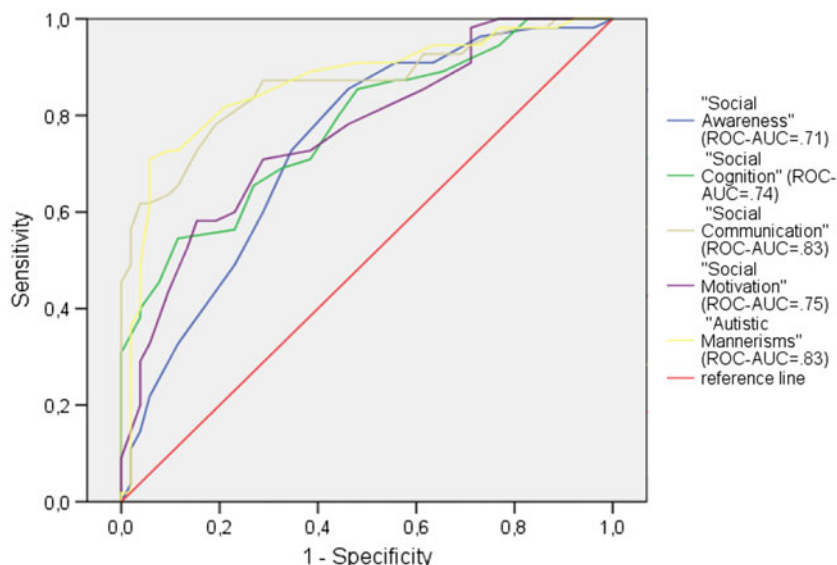


Fig. 2 Receiver Operating Characteristics (ROC) curve (AUC = 0.82, CI = 0.74–0.90) of the Social Responsiveness Scale (SRS) for Autism Spectrum Disorder (ASD) versus Oppositional Defiant Disorder (ODD)/Conduct Disorder (CD)

Fig. 3 Receiver Operating Characteristics (ROC) curve of the Social Responsiveness Scale (SRS) sub-scales for Autism Spectrum Disorder (ASD) versus Oppositional Defiant Disorder (ODD)/Conduct Disorder (CD)



Confidence Interval (CI) for sub-scales: *social awareness* (CI=.62-.81), *social cognition* (CI=.65-.83), *social communication* (CI=.75-.91), *social motivation* (CI=.65-.84), *autistic mannerisms* (CI=.75-.91)

(*hyperactivity, total score*) (Table 2). For these questionnaires, additional separate ROC-analyses were performed to explore their validity to differentiate between ASD and ODD/CD. The best ROC results were obtained for the SCQ (AUC = 0.84; 95 % CI = 0.77–0.92) and FBB-SSV (AUC = 0.19, 95 % CI = 0.11–0.28). A score of 0.5 would characterise a random result, whereas a result of AUC = 0.19 strongly indicates a non-ASD classification. The other questionnaires only reached ROC-AUC >0.19 and <0.70 (CBCL: *internalizing* AUC = 0.62, *externalizing* AUC = 0.28, *withdrawn* AUC = 0.68, *social problems* AUC = 0.67, *thought problems* AUC = 0.68, *delinquent behaviour* AUC = 0.29, *aggressive behaviour* AUC = 0.28; FBB-ANZ: *social anxiety* AUC = 0.69; FBB-ADHD: *hyperactivity* AUC = 0.37, *total score* AUC = 0.37). SCQ and FBB-SSV were entered together with the SRS into a binary logistic regression analysis to assess the improvement of the classification by the three report forms compared with the SRS alone. Cut-offs for the SCQ (total score 11) and SRS (total score 80) were chosen according to the best sensitivity and specificity for this population. For the FBB-SSV the manual criteria to classify ODD/CD were used as classification predictor. For classification, Stanine scores can be calculated with a mean of five and a standard deviation of two, scores ≥ 8 indicate a clinical diagnosis. Each of the three questionnaires improved the respective fit of the binary regression model (likelihood ratio tests: $p_{\text{all}} < 0.05$; Table 4). A combination of SRS, SCQ and FBB-SSV as independent predictors showed the best model accuracy ($\chi^2 = 64.07$, $p < 0.001$; $-2 \text{ Log Likelihood} = 58.39$) and explanatory value (Nagelkerkes $R^2 = 0.69$). Nagelkerkes R^2 summarizes how

Table 4 Binary logistic regression results comparing SRS alone and in combination with SCQ and FBB-SSV classifying ASD

	Model fit			Right classification (%)	Variables	<i>b</i> (SE)	Wald statistic (<i>df</i> = 3.84)	Odd's ratio	<i>p</i> value
	χ^2	-2 LL	<i>R</i> ²						
Model 1	27.89 (<i>p</i> < 0.000)	94.58	0.36	77.5	SRS	2.46 (0.51)	23.36	11.70	0.000
Model 2	39.56 (<i>p</i> < 0.000)	82.91	0.48	80.9	SRS	1.63 (0.58)	7.93	5.09	0.005
					SCQ	1.97 (0.59)	11.30	7.19	0.001
Model 3	64.07 (<i>p</i> < 0.000)	58.39	0.69	85.4	SRS	2.36 (0.86)	7.59	10.54	0.006
					SCQ	2.06 (0.76)	7.29	7.85	0.007
					FBB_SSV	-3.29 (0.85)	14.97	0.04	0.000

LL log likelihood, *R*² Nagelkerkes *R*², *b* regression coefficient, *SE* standard error, *SRS* social responsiveness scale, *SCQ* social communication questionnaire, *FBB_SSV* external assessment questionnaire for Oppositional Defiant Disorders and Conduct Disorder

much of the “variability” of the dependent variable is explained by the independent variables. The rate of correctly classified individuals increased from 77.5 % using the SRS only to 85.4 % using all three instruments with the respective cut-offs.

Convergent validity

Spearman correlations of the SRS total raw score and the SRS sub-scales with other autism rating scales or diagnostic measures were generally positive, but showed variability in size. The strongest correlation of the SRS total raw score was found with the SCQ total raw score (*r* = 0.55, 95 % CI = 0.28–0.74) in the ASD sample. Only moderate correlations with ADI-R algorithm scores were observed (*social interaction* *r* = 0.33, 95 % CI = 0.04–0.57; *social communication* *r* = 0.31, 95 % CI = 0.04–0.54., *stereotyped behaviour* *r* = 0.45, 95 % CI = 0.18–0.66).

Concurrent validity

For *n* = 155 individuals, SRS and CBCL data were available. The Spearman correlation between the total SRS and the total CBCL score was *r* = 0.78, 95 % CI 0.69–0.83 (*p* ≤ 0.001). All correlations between the SRS total score and the CBCL syndrome scales were also positive (Table 5). The strongest correlations were found for *social problems* (*r* = 0.81, 95 % CI = 0.75–0.85), *attention problems* (*r* = 0.80, 95 % CI = 0.73–0.85) and *social withdrawn* (*r* = 0.75, 95 % CI = 0.67–0.81).

Discussion

The SRS is a dimensional rating scale designed to measure autistic symptoms in the population. Clinically, the parent version is increasingly being used for screening purposes,

Table 5 Convergent and concurrent validities of the SRS with SCQ, ADI-R, and CBCL

	SRS	
	<i>N</i>	ρ (<i>p</i> value)
SCQ^a		
Total score	55	0.55 (≤0.001)
ADI-R^a		
Social domain	55	0.33 (≤0.05)
Communication domain	55	0.31 (≤0.05)
Stereotypes domain	55	0.45 (≤0.001)
CBCL		
Total score	155	0.78 (≤0.001)
Internalizing	155	0.77 (≤0.001)
Externalizing	155	0.61 (≤0.001)
Withdrawn	155	0.75 (≤0.001)
Somatic complaints	155	0.32 (≤0.001)
Anxious/depressed	155	0.70 (≤0.001)
Social symptoms	155	0.81 (≤0.001)
Thought problems	155	0.66 (≤0.001)
Attentional symptoms	155	0.80 (≤0.001)
Delinquent behaviour	155	0.55 (≤0.001)
Aggressive behaviour	155	0.57 (≤0.001)

^a For comparability convergent validity was calculated for the ASD group only

P spearman’s rho, *SRS* social responsiveness scale, *SCQ* social communication questionnaire, *ADI-R* autism diagnostic interview-revised, *CBCL* child behaviour checklist, *ASD* autism spectrum disorder

especially in high functioning individuals suspected to have ASD. When used clinically, validity needs to be not only established comparing ASD and TD, but also ASD and children and adolescents with other psychiatric diagnoses. The present study is the first to specifically study diagnostic validity in ASD and ODD/CD patients. This is especially relevant, as high SRS scores were also described for ODD/CD, and not only for ASD [7, 10]++, and

children affected with either disorder exhibit difficulties in social interaction and reciprocal behaviour. Thus, the primary aim of the study was to evaluate the validity of the SRS to differentiate between ASD and ODD/CD.

As hypothesized we observed highest SRS total scores in children with ASD, followed by ODD/CD and TD. These data are in keeping with previous studies [7, 10, 16] with the difference that our findings show somewhat lower scores for all three samples [7]. Because of the greater overlap ODD/CD and ASD symptoms in the high functioning range, we selected our sample with $IQ \geq 70$. Autistic children with comorbid intellectual delay often show higher rates of general psychopathology, which may explain higher SRS total scores in the US and German standardization studies, which also included ASD children with low IQ [7, 14]. Another reason could be that a substantial part of the US standardization sample consisted of twins [23]. Corresponding to previous studies, no differences between ASD subgroups were observed [7, 13]. This supports the DSM-5 concept of Autistic Disorder as spectrum disorders instead of separate categories. A lack of differentiation between ASD subgroups by the SRS was also reported by Constantino et al. [17]. Similarly, ODD and CD did not differ with regard to the total SRS score, supporting the concept of analysing both disorders together.

As also shown in US and German studies [7, 14], the SRS total score distinguished excellently between ASD and TD ($AUC = 1.0$). In contrast, ROC-AUC was significantly lower when ASD and ODD/CD were compared. The respective sensitivity (0.76) and specificity (0.82) are in line with previous studies comparing ASD with a mixed group of other mental disorders (0.73, 0.81 in Bölte and Poustka [7]; 0.70, 0.90 in Constantino and Gruber [14]; 0.66, 0.89 in Wang et al. [40]). These results prove our second hypothesis of a lower validity to differentiate ASD versus ODD/CD than ASD versus TD, which is indicated by non-overlapping 95 % confidence intervals of the ROC-AUCs in the present study. When differentiating ASD and ODD/CD with the cut-off value = 80 as indicated by ROC-AUC, only 76 % of ASD were correctly classified as positives, what means simultaneously that 24 % children and adolescents with ASD were not identified by the SRS when compared with ODD/CD at the respective cut-off value. On the other hand, 82 % of ODD/CD were correctly classified as non-ASD, but 18 % were wrongly allocated to the ASD group as false positives. Despite the fact that sensitivity is one of the most important and critical aspects of a screening instrument, neither the US-, German or Taiwanese standardization studies nor our data showed sufficient sensitivity of the SRS total score to differentiate between ASD and other child psychiatric disorders. Also further research groups, like Charman et al. [13] reported a

substantial lower specificity (0.67) comparing children with educational needs with and without ASD and an even more reduced specificity in subsamples with low IQ (0.57) or additional behaviour problems (0.41). Discriminant validity of the sub-scales in our study was better for *autistic mannerisms* and *social communication*. Both scales are also the two scales with the best internal consistency in the German standardization study ($\alpha = 0.90$, $\alpha = 0.92$) [7]. Sensitivity and specificity are determined by the cut-off score. The higher the cut-off score, the greater the specificity, but simultaneously the sensitivity decreases as the most important aspect of a screening tool and vice versa. To better translate the meaning of the ROC into everyday diagnostics of clinicians and researchers, we additionally calculated the positive and negative predictive values, applying the “best” cut-off of 80 at maximum sensitivity and specificity to differentiate between ASD and ODD/CD. The probability to correctly classify ASD by an SRS score ≥ 80 was 81 %. A negative test (SRS score < 80) gave a 77 % probability for correct classification as non-ASD respective ODD/CD. These results underline the difficulties of the SRS to differentiate reliably between ASD and ODD/CD.

Therefore, in this study we explored if a combination of disorder-specific questionnaires for ASD and ODD/CD would improve correct classification. This has a direct clinical impact, as the rate of wrongly classified individuals for both disorders can be reduced by such an approach. The combination of different screening and disorder-specific questionnaires increased the correct classification from 77.5 to 85.4 %. The model fit increased by adding SCQ and FBB-SSV, and every single questionnaire generated a significant contribution to the regression model confirming the respective model accuracy. Because of the improved specificity of the diagnostic process by using additional information, Corsello et al. [18] also suggested a multistage assessment beginning with the SCQ, followed by ADOS and ADI-R in diagnosing ASD. Here, we propose a different and more economic approach by combining information of three parent rating questionnaires before confirming an ASD diagnosis with ADI-R and ADOS.

We also explored the validity of the other questionnaires used in this study to assess their ability to differentiate between both clinical groups. Higher scores for ODD/CD in the CBCL scales *externalizing*, *delinquent* and *aggressive behaviour* and on both scales of the FBB-SSV (*conduct disorder*, *oppositional defiant disorder*) are perfectly in line with the diagnostic criteria [4, 42]. The higher scores for ASD compared with ODD/CD in the CBCL scales of *internalizing behaviour*, *withdrawn*, *social problems* and *thought problems* match well with ASD symptom patterns [4, 42], describing additional psychopathologic symptom variety besides the core phenomenology [9].

Despite these mean differences between groups, most of the FBB-ANZ, FBB-ADHD and some of the CBCL scores including the *total score* could not differentiate satisfyingly between clinical groups, indicating a strong overlap of these additional symptoms in both disorders. This is also reflected by the psychiatric comorbidity pattern and rate as assessed by the Kinder-DIPS in this study. This finding underlines the necessity of diagnostic accuracy studies in addition to describe mean differences on the group level. For clinical ASD screening purposes, therefore, the use of the disorder-specific questionnaires SRS, SCQ, and FBB-SSV clearly is recommended above the other scales. The other scales may be used to alert the clinician to the presence of comorbid disorders, which are high in ASD and ODD/CD and are of strong clinical relevance. International classification systems currently do not allow a large number of comorbid disorders in children with ASD. Not taking the ICD-10 or DSM-IV-TR exclusion criteria for ADHD or social phobia into account, the present study showed that more than half of the children with ASD exhibited comorbid symptoms fulfilling diagnostic criteria. By the same approach, Simonoff et al. [38] even identified a 70 % comorbidity rate in a population-based sample of children with ASD. Similarly, population-based studies in children with ODD/CD showed a 50 % comorbidity rate with ADHD, comparable to our sample [30].

Despite the inability to differentiate ASD from ODD/CD by CBCL-derived scales, concurrent validity with the SRS total score was high for the CBCL. Comparable to previous findings on CBCL scores in autism [9, 10] and fitting into the symptom pattern of ASD, our analysis on concurrent validity demonstrated the highest correlation with the CBCL sub-scales *social problems*, *attention problems* and *withdrawn*. Similarly to previous studies [7] the lowest correlation was found for *somatic complaints*. This high concurrent validity is in line with the notion that the SRS also is a measure of general psychopathology (e.g. [24]). Consistent with previous findings by Bölte and Poustka [7] or Charman et al. [13] convergent validity with the ADI-R was lower than reported for the US original [15, 17]. In contrast to Constantino et al. [17] our sample did not include other psychiatric disorders besides ASD, what may explain the lower correlation due to lower sample variance. Correlation between SRS and SCQ was in the medium range ($r = 0.55$) and similar to the findings of Bölte and Poustka ($r = 0.58$) [7]. This shows that SRS and SCQ describe an overlapping phenotype, but still add rating scale-specific information in the screening and diagnostic process. This fits well into the results of our logistic regression procedure showing that a combination of both questionnaires resulted in a higher rate of correctly classified individuals than the SRS alone.

An important strength of our study is the high internal validity, given by the matching procedure of the three groups, and the standardized and comprehensive clinical assessment. Gold standard diagnostic systems were used for ASD and ODD/CD, and comorbid diagnoses were obtained by an extensive parent interview. IQ data and parent questionnaire-based information about general psychopathology were collected, also in TD. Evaluation of the questionnaires took place without any knowledge of the children's diagnoses. Further, the calculation of confidence intervals for the instrument parameter estimates reflects the statistical strengths of the findings. These are quality aspects, which have not met by previous studies on the SRS (e.g. [10, 12]). In addition, the rate of comorbid psychiatric disorders of both clinical samples fits very nicely with clinical and epidemiological data of both disorders, with highest rates of ADHD in both groups [38]. The same pattern was also observed for the CBCL scales. This underlines the external validity of the study results and the representativeness of the described sample. Going beyond previous studies, in addition to evaluating the differential validity of the SRS, we aimed at improving diagnostic test accuracy by a combination of different disorder-specific questionnaires in addition to the SRS. This combined approach is economic and may be used in clinical practice. Before generally recommending this approach for clinical use, however, the findings should be replicated in an independent sample.

A shortcoming of our study is the fact that we used a combination of questionnaires implementing cut-off values derived from the same sample. Furthermore, the children's diagnoses were previously known by parents filling out the report forms. Thus, rater-bias cannot totally be excluded. This also may have resulted in an overestimation of diagnostic validity that should be taken into account in study interpretation. Convergent validity may be additionally exaggerated because all questionnaires were completed by parents, not by teachers or the participants themselves. Finally, we cannot exclude a recruitment bias in the TD group that may possibly not be representative of the general population.

Altogether, there is a pressing need for future studies. Ideally, a replication, especially of the binary regression analyses, should be done in a prospectively collected clinical population with unknown previous diagnoses to avoid rater-bias. Further research is also encouraged to compare the diagnostic criteria that differentiate disorders with difficulties in social interaction and communication better. Additional studies on cut-off scores [18] dependent on differential diagnoses (e.g. social phobia) and screening purposes would be helpful for clinicians. Another interesting approach would be an examination of specific item sets of the SRS to determine whether a different

combination or reduced (increased) number of items will distinguish the disorders in a more accurate way.

In conclusion, the current study replicated the excellent validity of the SRS to differentiate between ASD and TD, but alerts to the possibility of false-positive ASD diagnoses in children with ODD/CD and false-negative diagnoses in children with ASD by the SRS, depending on the respective cut-off values. AUC was not sufficient to support the use of the SRS as the only screening instrument differentiating ASD and ODD/CD, but a combination of SRS and SCQ with an ODD/CD specific questionnaire based on DSM-IV TR (the FBB-SSV) can be recommended for clinical practice. When aiming at implementing cost- and time-effective instruments in research and clinical practice, the SRS can be helpful when the described limitations are considered as well. Those who are interested in using the SRS should consider adjusting cut-off scores due to their purpose [18] and consider the overlapping symptoms between disorders. Awareness that social impairment is not an exclusive symptom of ASD is indispensable. In order to obtain valid screening procedures, a multiple instrument approach is certainly necessary [18, 39]. Finally, this study illustrates the importance to keep in mind that a score in a screening instrument cannot replace an extensive diagnostic assessment. For detailed ASD diagnosis a combination of parental interview, structured observation of the child's behaviour and considering comorbid and differential psychiatric disorders is necessary to ensure that the child will receive the respective targeted therapy.

Acknowledgments The authors gratefully acknowledge all families and their children who took part in this study. The authors also thank Laura Mojica and Eva Westerwald for their support in data collection. The study was supported by a grant of the Medical Faculty of the JW Goethe University Frankfurt am Main (Heinrich and Fritz Riese foundation)

Conflict of interest The authors declare that they have no conflict of interest.

References

- Achenbach TM (1998) Elternfragebogen über das Verhalten von Kindern und Jugendlichen; Deutsche Bearbeitung der Child Behaviour Checklist (CBCL/4-18). Arbeitsgruppe Kinder-, Jugend- und Familiendiagnostik, Köln
- Adams C, Green J, Gilchrist A, Cox A (2002) Conversational behaviour of children with Asperger syndrome and conduct disorder. *J Child Psychol Psychiatry* 43(5):679–690
- Adornetto C, In-Albon T, Schneider S (2008) Diagnostik im Kindes- und Jugendalter anhand strukturierter Interviews: Anwendung und Durchführung des Kinder-Dips. *Klinische Diagnostik und Evaluation* 1(4):363–377
- American Psychiatric Association (2000) Diagnostic and statistical manual of mental disorders, 4th ed., text revision. American Psychiatric Association, Washington, DC
- Aster M, Neubauer A, Horn R (2006) Wechsler Intelligenztest für Erwachsene (WIE). Harcourt Test Services, Frankfurt a. Main
- Bölte S, Poustka F (2006) Fragebogen zur Sozialen Kommunikation—Autismus Screening (FSK). Deutsche Fassung des Social Communication Questionnaire (SCQ). Hans Huber, Hogrefe AG, Bern
- Bölte S, Poustka F (2008) Skala zur Erfassung Sozialer Reaktivität (SRS). Hans Huber, Hogrefe AG, Bern
- Bölte S, Crecelius K, Poustka F (2000) Der Fragebogen über Verhalten und soziale Kommunikation: psychometrische Eigenschaften eines Autismus-Screening-Instruments für Forschung und Praxis. *Diagnostica* 46:149–155
- Bölte S, Dickhut H, Poustka F (1999) Patterns of parent-reported problems indicative in autism. *Psychopathology* 32:94–98
- Bölte S, Poustka F, Constantino JN (2008) Assessing autistic traits: cross-cultural validation of the social responsiveness scale (SRS). *Autism Res* 1:354–363
- Bölte S, Rühl D, Schmötzer G, Poustka F (2006) Diagnostisches Interview für Autismus—Revidiert. Deutsche Fassung des Autism Diagnostic Interview—Revised von Michael Rutter, Ann LeCouteur und Catherine Lord, 1. Auflage. Hans Huber, Hogrefe AG, Bern
- Bölte S, Westerwald E, Holtmann M, Freitag C, Poustka F (2011) Autistic traits and autism spectrum disorders: the validity of two measures presuming a continuum of social communication skills. *J Autism Dev Disord* 42(1):66–72
- Charman T, Baird G, Siminoff E, Loucas T, Chandler S, Meldrum D, Pickles A (2007) Efficacy of three screening instruments in the identification of autistic-spectrum disorders. *Br J Psychiatry* 191:554–559
- Constantino JN, Gruber CP (2005) Social responsiveness scale. Western Psychological Services, Los Angeles
- Constantino JN, Hudziak JJ, Todd RD (2003) Deficits in reciprocal social behaviour in male twins: evidence for a genetically independent domain of psychopathology. *J Am Acad Child Adolesc Psychiatry* 42(4):458–467
- Constantino JN, Przybeck T, Friesen D, Todd R (2000) Reciprocal social behaviour in children with and without pervasive developmental disorders. *Dev Behav Pediatr* 21(1):2–11
- Constantino J, Davis S, Todd R, Schindler M, Cross M, Brophy S et al (2003) Validation of a brief measure of autistic traits: comparison of the social responsiveness scale with the autism diagnostic interview-revised. *J Autism Dev Disord* 33(4):427–433
- Corsello C, Hus V, Pickles A, Risi S, Cook EH, Leventhal BL et al (2007) Between a ROC and a hard place: decision making and making decisions about using the SCQ. *J Child Psychol Psychiatry* 48:932–940
- Döpfner M, Görtz-Dorten A, Lehmkuhl G (2008) DISYPS-II: Diagnostik System für psychische Störungen nach ICD-10 und DSM-IV für Kinder und Jugendliche-II. Huber, Göttingen
- Döpfner M, Schmeck K, Berner W (1994) Handbuch: Elternfragebogen über das Verhalten von Kindern und Jugendlichen. Forschungsergebnisse zur deutschen Fassung der Child Behaviour Checklist (CBCL/4-18). Arbeitsgruppe Kinder-, Jugend- und Familiendiagnostik, Köln
- Eaves L, Rutter M, Silberg J, Shillady L, Maes H, Pickles A (2000) Genetic and environmental causes of covariation in interview assessments of disruptive behaviour in child and adolescent twins. *Behav Genet* 30(4):321–334
- Grzadzinski R, Di Martino A, Brady E, Mairena MA, O'Neale M, Petkova E, Lord C, Castellanos FX (2011) Examining autistic traits in children with ADHD: does the autism spectrum extend to ADHD? *J Autism Dev Disord* 41:1178–1191
- Ho A, Todd RD, Constantino JN (2005) Brief report: autistic traits in twins vs non-twins—a preliminary study. *J Autism Dev Disord* 35:129–133

24. Hus V, Bishop S, Gotham K, Huerta M, Lord C (2013) Factors influencing scores on the social responsiveness scale. *J Child Psychol Psychiatry* 54:216–224
25. Kanne SM, Abbacchi AM, Constantino JN (2009) Multi-informant ratings of psychiatric symptom severity in children with autism spectrum disorders: the importance of environmental context. *J Autism Dev Disord* 39:856–864
26. Lecavalier L (2006) Behavioural and emotional problems in young people with pervasive developmental disorders: relative prevalence, effects of subject characteristics, and empirical classification. *J Autism Dev Disord* 36:1101–1114
27. Lecavalier L, Gadow KD, DeVincenzi CJ, Edwards MC (2009) Validation of DSM-IV model of psychiatric syndromes in children with autism spectrum disorders. *J Autism Dev Disord* 39:278–289
28. Lord C, Rutter M, DiLavore P, Risi S (2001) Autism diagnostic observation schedule (ADOS). Western Psychological Services, Los Angeles
29. Mandell DS, Ittenbach RF, Levy SE, Pinto-Martin JA (2007) Disparities in diagnoses received prior to a diagnosis of autism spectrum disorder. *J Autism Dev Disord* 37(9):1795–1802
30. Peterman F (2010) Störung des Sozialverhaltens. In: Mattejat F (ed) *Lehrbuch der Psychotherapie*, 4th edn. CIP Medien, München
31. Petermann F, Petermann U (2010) *Hamburg-Wechsler-Intelligenztest für Kinder-IV*. Hans Huber, Bern
32. Risi S, Lord C, Gotham K, Corsello C, Chrysler C, Szatmari P, Cook E, Bennett L, Leventhal MD, Pickles A (2006) Combining information of multiple sources in the diagnosis of autism spectrum disorders. *J Am Acad Child Adolesc Psychiatry* 45(9):1094–1103
33. Rühl D, Bölte S, Feineis-Matthews S, Poustka F (2004) Diagnostische Beobachtungsskala für Autistische Störungen. Deutsche Fassung der Autism Diagnostic Observation Schedule. Hans Huber, Hogrefe AG, Bern
34. Rutter M, Bailey A, Lord C (2003) *Social communication questionnaire (SCQ)*. Western Psychological Services, Los Angeles
35. Rutter M, LeCouteur A, Lord C (2003) *Autism diagnostic interview-revised (ADI-R)*. Western Psychological Services, Los Angeles
36. Schneider S, Unnewehr S, Margraf J (2009) *Diagnostisches interview bei psychischen Störungen im Kindes- und Jugendalter*, 2nd edn. Springer Verlag, Heidelberg
37. Schwenck C, Mergenthaler J, Keller K, Zech J, Salehi S, Taurines R, Romanos M, Schecklmann M, Schneider W, Warnke A, Freitag CM (2012) Empathy in children with autism and conduct disorder: group-specific profiles and developmental aspects. *J Child Psychol Psychiatry* 53(6):651–659
38. Simonoff E, Pickles A, Charman T, Chandler S, Loucas T, Baird G (2008) Psychiatric disorders in children with autism spectrum disorders: prevalence, comorbidity and associated factors in a population-derived sample. *J Am Acad Child Adolesc Psychiatry* 47(8):568–572
39. Towbin KE, Pradella A, Gorrindo T, Pine DS, Leibenluft E (2005) Autism spectrum traits in children with mood and anxiety disorders. *J Child Adolesc Psychopharmacol* 15(3):452–464
40. Wang J, Lee L-C, Chen Y-S, Hsu J-W (2012) Assessing autistic traits in a Taiwan preschool population: cross-cultural validation of the social responsiveness scale (SRS). *J Autism Dev Disord* 42(11):2450–2459
41. Weiss RH (2006) *CFT 20-R: Grundintelligenz Skala 2-Revision*. Hogrefe, Göttingen
42. World Health Organisation (1992) *International statistical classification of diseases and related health problems Rev. 10*. World Health Organisation, Geneva