



Screening for Autism Spectrum Disorder with the Achenbach System of Empirically Based Assessment Scales

Anne Deckers¹ · Peter Muris^{1,2} · Jeffrey Roelofs¹

Published online: 31 July 2019
© The Author(s) 2019

Abstract

Integrating a screening possibility for Autism Spectrum Disorder (ASD) within the widely applied Achenbach System of Empirically Based Assessment (ASEBA) scales could be of great value for daily clinical practice. The present study explored the ability of the school-aged Child Behavior Checklist (CBCL) and the Teacher's Report Form (TRF) to screen for ASD in children and adolescents (aged 6 to 18 years) within a mixed clinically referred sample. Different screening variants were compared for the CBCL, TRF, and the combination of CBCL and TRF: the separate withdrawn/depressed, social problems, and thought problems syndrome scales; combinations of these syndrome scales; and special ASD scales composed of relevant individual items. Analyses were performed for: youth with a DSM-IV-based clinical diagnosis; youth for which the clinical DSM-IV diagnosis was confirmed by a standardized assessment (the Autism Diagnostic Interview-Revised); and youth with a DSM-IV-based clinical diagnosis of ASD that also met the DSM-5 criteria. The results clearly demonstrated that the special ASEBA-based scales – in particular when completed by the parents – were most predictive of ASD. The results also indicate that following the initial screen with these ASEBA scales, further thorough diagnostic assessment is necessary to definitively establish whether young people really suffer from ASD.

Keywords Autism spectrum disorders · Achenbach system of empirically based assessment (ASEBA) · Children and adolescents · Screening

Autism spectrum disorder (ASD) is a pervasive developmental disorder characterized by persistent deficits in social interaction and communication in combination with restricted, obsessive, or repetitive patterns of behavior (American Psychiatric Association 2000, 2013). Due to the polythetic criteria as used in psychiatric classification systems, ASD is quite heterogeneous in terms of symptom composition and severity. Youth with severe forms of ASD as described in the DSM-5 or the classical Autistic Disorder according to the DSM-IV can be identified rather well. However, children and adolescents with milder forms of ASD like Pervasive Developmental Disorder-Not Otherwise Specified (PDD-NOS) or high-functioning

children with ASD are more difficult to recognize. In fact, high-functioning youth with ASD are often diagnosed relatively late, that is in elementary or even secondary school (Daniels and Mandell 2014), or the ASD may not be recognized at all (Kim et al. 2011). When these young people are referred to mental health care, they often display a wide range of secondary complaints like anxiety, depression, and behavioral problems (Simonoff et al. 2008). They are often 'diagnostic puzzles' for clinicians and thorough diagnostic assessment is required. The Autism Diagnostic Interview-Revised (ADI-R; Rutter et al. 2003) and Autism Diagnostic Observation Scale (ADOS-2; Lord et al. 2012) are helpful standardized instruments to support the diagnostic process, but the administration of these measures is quite time-consuming, labor-intensive, and thus relatively expensive (Howlin and Asgharian 1999). Easy-to-administer screening instruments that indicate the possible presence of ASD in school-aged youth would be welcome as they enable clinicians to quickly identify those at risk for ASD who require an in-depth examination and to prevent others from redundant or irrelevant diagnostic procedures.

✉ Anne Deckers
anne.deckers@maastrichtuniversity.nl

¹ Youz (Parnassia Group) and Department of Clinical Psychological Science, Maastricht University, Postbox 616, 6200, MD Maastricht, The Netherlands

² Departement Sielkunde, University of Stellenbosch, Stellenbosch, South Africa

Several instruments have been developed that can be used to screen for ASD in school-aged children and adolescents, including the Social Responsiveness Scale (SRS; Constantino and Gruber 2005), the Social Communication Questionnaire (SCQ; Berument et al. 1999), the Children's Communication Checklist (CCC; Bishop 1998), and the Autism-Spectrum Quotient (Auyeung et al. 2008; Baron-Cohen et al. 2006). Despite the fact that these ASD screening instruments are psychometrically sound and available, they are not regularly applied in daily clinical practice – probably because there is little room during the intake procedure for administering extra assessment instruments with a special focus on only one type of psychopathology. However, in many countries (including The Netherlands) clinics routinely apply the Achenbach System of Empirically Based Assessment (ASEBA; original version: Achenbach 1991; revised version: Achenbach and Rescorla 2001) scales as a broad screen to get an initial impression of the behavioral, emotional, and social problems (as well as competencies) of clinically referred children and adolescents. The ASEBA scales make use of multiple informants: Parents complete the Child Behavior Checklist (CBCL), teachers the Teacher Report Form (TRF), while children themselves from 11 years onwards fill in the Youth Self Report (YSR). Interestingly, it has been suggested that the ASEBA scales might also be valuable for the screening of ASD. Routine-wise scoring by means of these scales might increase clinicians' awareness of ASD and might lead to better detection of this disorder and indication of appropriate interventions. It is good to note that the preschool version of the Child Behavior Checklist (CBCL 1.5–5; Achenbach and Rescorla 2000) already contains a DSM-oriented scale of pervasive developmental problems that can be effectively used to screen for ASD (e.g., Havdahl et al. 2016; Myers et al. 2014; Muratori et al. 2011; Sikora et al. 2008), but in the school-aged version of the Achenbach scales (for children and adolescents of 6–18 years) a standard subscale for assessing the symptomatology of ASD has not been included. However, both the original and the revised ASEBA scales contain various items (e.g., “Strange behavior”; “Doesn't get along with other kids”; “Would rather be alone than with others”) and syndrome scales (i.e., depressed-withdrawn, thought problems, social problems) that are highly relevant for ASD. Thus, it is not surprising that researchers have begun to evaluate individual syndrome scales, combinations of these scales, and specific sets of items of the ASEBA scales as potential screeners to detect children and adolescents that might be suffering from this type of psychopathology.

Most studies have relied on the parent version of the Achenbach scales, the CBCL, and made use of the existing syndrome scales to explore whether these are statistically different for youth with and without ASD. A first attempt was made by Bölte and colleagues (Bölte et al. 1999) who compared school-aged CBCL subscale scores of youths aged 4 to

18 years who were diagnosed with autistic disorder and a control group of youth with a mix of other psychiatric problems. It was found that the ASD youth scored higher on social problems, thought problems, and attention problems but lower on somatic complaints as compared to youth in the clinical control group. Adopting a similar approach in a small Brazilian sample of clinic-referred children aged 4 to 11 years, Duarte et al. (2003) noted that only the thought problems subscale of the CBCL differentiated between youth with ASD and youth with other psychiatric disorders. In a further study by Hoffmann and colleagues (Hoffmann et al. 2016), the screening utility of the CBCL syndrome scales was explored in an outpatient sample of high-functioning (IQ > 70) 4- to 18-year-old children and adolescents with ASD. The researchers made comparisons between an ASD only group, an ASD plus comorbid ADHD group, and control groups of children and adolescents with only ADHD or with internalizing disorders only. Only the social problems subscale was found to differentiate between both ASD groups and the control groups.

Biederman et al. (2010) took the research on this topic one step further by not only examining whether individual CBCL syndrome scales differentiated youth with (high-functioning) ASD from clinically referred youth without this condition (the mean age in both groups was on average 11 years), but also by looking at the combination of scales. Logistic regression analyses revealed that the withdrawn/depressed, social problems, and thought problems syndrome scales were significant predictors of an ASD diagnosis, but Biederman et al. (2010) also demonstrated that an ASD profile consisting of the combination of these three subscales had the best potential to discriminate between youth with and without ASD. In an attempt to replicate these results, Havdahl et al. (2016) found that children with ASD aged 6–13 years scored significantly higher on the CBCL syndrome scales withdrawn/depressed, social problems, and thought problems than children in a mixed clinical control group. However, these authors also critically noted that the capacity of these subscales to predict the diagnosis of ASD was relatively weak. Within this sample, the combination of the withdrawn/depressed and thought problems scales (which the researchers named the WTP profile) turned out to have, relatively speaking, the best discriminative power.

Using a clinical sample of 4- to 18-year-old children and adolescents from Singapore, Ooi and colleagues (Ooi et al. 2011) also examined the utility of the school-aged CBCL as a screener for ASD. In line with the previous findings, the withdrawn/depressed and thought problems syndrome scales differentiated between youth with ASD and youth in other clinical comparison groups (i.e., attention-deficit/hyperactivity disorder (ADHD) – inattentive type, ADHD – hyperactive-impulsive or combined subtype, and a referred but undiagnosed group). The social problems subscale had less

discriminative power and only differentiated youth with ASD group from youth with ADHD – inattentive type and the referred but undiagnosed group. Interestingly, the researchers also construed a special ASD scale that only contained individual CBCL items that discriminated between the ASD group and the comparison groups. The results showed that this ASD scale had better discriminative power than each of the three original syndrome scales.

A final study that seems relevant to discuss within this context was conducted by So et al. (2013) who also construed a special ASD scale based on ASEBA items for which scores statistically differed between youth with ASD and clinically referred youth with other psychiatric conditions and typically developing children and adolescents. Strong points of this study were that (a) the selection of items and the subsequent testing of the screening potential of the scale were performed in different samples (ages 6 to 18 years); and (b) that not only parents were used as informant (by completing the CBCL), but also the teachers (who filled in the TRF). Teacher information might make a valuable contribution to the detection of children and adolescents at risk for ASD as they have the opportunity to observe youth's functioning within the school setting and especially during interaction with peers. Thus, So et al. (2013) investigated the discriminative power of a CBCL-based, a TRF-based, and a combined CBCL/TRF-based ASD scale. The results demonstrated that both the CBCL- and the TRF-based ASD scales were able to discriminate between youth with ASD and youth with internalizing and externalizing disorders, referred children and adolescents without a diagnosis, and the total clinical control group. In addition, results showed that combining the parent and teacher information (the CBCL/TRF-based ASD scale) improved the chance of correctly identifying children and adolescents with ASD.

Altogether, research exploring the school-aged ASEBA scales as a potential screener for ASD seems to indicate that: (1) although the original syndrome scales can be used for this purpose, there are indications that it is better to construe and employ a special scale consisting of a number of selected items (Ooi et al. 2011; So et al. 2013); and (2) screening based on multiple informants (i.e., parents and teachers) may have better predictive value than screening based on only one informant (So et al. 2013), which is also well in line with the general recommendation regarding the assessment of child and adolescent psychopathology (De Los Reyes and Kazdin 2005). Given the potential benefits and utility of an ASD screener that is incorporated in a widely used assessment instrument such as the ASEBA scales, and given the small number of studies conducted so far, it seemed worthwhile to further investigate the utility of the CBCL and TRF as a screen for ASD.

With this in mind, the present investigation was designed to examine how the school-aged ASEBA scales can best be used for this purpose. In a mixed sample of clinically referred

children and adolescents, we compared various ways of using this instrument for the identification of youth with ASD. A number of screening variants were evaluated, which relied on: (1) the separate withdrawn/depressed, social problems, and thought problems syndrome scales; (2) the ASD profile of Biederman et al. (2010), which combines these three syndrome scales; (3) the WTP scale as proposed by Havdahl et al. (2016), which only combines the withdrawn/depressed and thought problems syndrome scales; (4) the special ASD scale as construed by Ooi et al. (2011); and (5) the special ASD scale as developed by So et al. (2013; see Table 1 for an overview of the scales and the corresponding items). This research builds on previous studies but also extends this work in two important ways: (1) a consequent multi-informant approach was taken, implying that we tested each screening method not only relying on parent-report (CBCL) but also on teacher-report (TRF) data; (2) the capacity to discriminate between youth with ASD from other clinically referred youth may also critically depend on how ASD is defined. Currently, we employ the diagnostic criteria as defined in the DSM-5 (American Psychiatric Association 2013), whereas all the previous studies on the screening potential of the school-aged ASEBA scales used DSM-IV-based criteria to classify ASD. Further, although there is considerable consensus that standardized assessments such as the ADI-R and the ADOS-2 increase the diagnostic accuracy of ASD (Falkmer et al. 2013), their use in clinical practice as well as research is not widespread (Lord 2010) and indeed the earlier studies evaluating the ASEBA scales as a screen often did not make use of a standardized instrument to confirm the diagnosis of ASD. Therefore, in order to make a good comparison with the previous empirical work, we first of all tested the potential of the ASEBA instrument to screen for youth with a DSM-IV-based clinical diagnosis of ASD without the confirmation of a standardized assessment instrument. Subsequently, we examined the screening potential of the ASEBA scales for DSM-IV and DSM-5 diagnoses of ASD that were also confirmed by the ADI-R (Rutter et al. 2003). It was hypothesized that, irrespective of how the ASD diagnosis was defined, (1) the special ASD scales construed by Ooi et al. (2011) and So et al. (2013) would be better in screening for ASD than the existing syndrome scales; and (2) screening based on multiple informants (i.e., the combination of CBCL and TRF data) has better predictive value than screening based on only one informant.

Method

Participants

The total sample, which in the further course of this paper will be referred to as the *clinical DSM-IV diagnosis group*, consisted of 132 participants of whom: 75 were diagnosed

Table 1 An overview of the ASEBA, CBCL, and TRF scales/items used to screen for ASD

ASD profile (Biederman et al. 2010)				
WTP scale (Havdahl et al. 2016)				
Withdrawn/depressed	Thought problems	Social problems	ASD scale (Ooi et al. 2011)	ASD scale (So et al. 2013)
There is little that he/she enjoys	Cannot get his/her mind off certain thoughts; obsessions	Clings to adults or too dependent	Acts too young for his/her age	Acts too young for his/her age
Would rather be alone than with others	Deliberately harms self or attempts suicide	Complains of loneliness	Doesn't get along with other kids	Cannot get his/her mind off certain thoughts; obsessions
Refuses to talk	Hears sounds or voices that are not there	Doesn't get along with other kids	Fears certain animal, situations	Daydreams or gets lost in his/her thoughts
Secretive, keeps things to self	Nervous movements or twitching	Easily jealous	Would rather be alone than with others	Would rather be alone than with others
Too shy or timid	Picks nose, skin, or other parts of body	Feels others are out to get him/her	Nervous movements or twitching	Poorly coordinated or clumsy
Underactive, slow moving, or lacks energy	Repeats certain acts over and over; compulsions	Gets hurt a lot	Repeats certain acts over and over; compulsions	Repeats certain acts over and over; compulsions
Unhappy, sad, or depressed	Sees things that are not there	Gets teased a lot	Speech problem	Speech problem
Withdrawn, doesn't get involved with others	Stores up too many things he/she does not need	Not liked by other kids	Strange behavior	Stares blankly
	Strange behavior	Poorly coordinated or clumsy	Withdrawn, doesn't get involved with others	Strange behavior
	Strange ideas	Prefers being with younger kids		Withdrawn, doesn't get involved with others
	<i>CBCL only:</i>	Speech problem		
	Plays with own sex parts in public			
	Plays with own sex parts too much			
	Sleeps less than most kids			
	Talks or walks in sleep			
	Trouble sleeping			

ASEBA Achenbach Scales of Empirically Based Assessment, CBCL Child Behavior Checklist, TRF Teacher Report Form, ASD Autism Spectrum Disorder, WTP Withdrawn/depressed and Thought problems

with ASD, and 57 were not diagnosed with ASD but with other DSM-IV diagnoses (clinical control group; see left columns of Table 2).¹ In the ASD group, 71% was diagnosed with PDD-NOS, 19% with Asperger's Disorder, and 11% with Autistic Disorder. The majority of the participants in the clinical control group had ADHD of the combined (47%) or inattentive (23%) subtype as primary diagnosis. In both groups, comparable comorbidity rates were found [$\chi^2(1) = 1.50, p = .22$]. In the ASD group, 84% of the children were diagnosed with at least one comorbid diagnosis and 43% even had 2 or more comorbid diagnoses. In the clinical control group, these percentages were 75% and 39%, respectively. The sample mainly contained participants with a Caucasian ethnic background (>90%). There were more boys than girls (79% versus 21%) and the mean age was 11.63 years ($SD = 3.32$), with a range of 6 to 18 years. IQ scores were only available for 67.4% of the participants and varied between 71 and 135 with an average IQ of 100.30 ($SD = 14.58$). There were no significant differences between the ASD and clinical control groups with regard to age [$F(1, 132) = .14, p = .71$], gender distribution [$\chi^2(1) = 2.82, p = .09$], and IQ [$F(1, 89) = 2.95, p = .09$].

¹ Due to missing variables, the exact n 's for the ASD and control groups slightly vary for each analysis. When <10% of the items of a scale was missing, values were imputed with the mean score of that participant on that particular scale. In case of more missing items, listwise deletion per analysis was employed.

To test the screening potential of the ASEBA scales under more stringent diagnostic conditions, we also made comparisons between ASD and clinical control participants for whom the diagnostic status (i.e., ASD or clinical control) had been confirmed by the ADI-R. Participants for whom there was no consensus between the clinical DSM diagnosis and the outcome of the ADI-R were discarded. In case the ADI-R was scored according to the DSM-IV criteria, 35 participants were removed, leaving 48 participants in the *ADI-R supported DSM-IV diagnosis ASD group* versus 49 participants in the clinical control group. When scoring the ADI-R by means of the DSM-5 criteria, 29 participants were discarded, leaving 57 participants in the *ADI-R supported DSM-5 diagnosis ASD group* versus 46 participants in the clinical control group. Irrespective of the DSM criteria that were used, the ADI-R supported ASD and clinical control groups did not differ in terms of age, gender distribution, and IQ (see Table 2).

Procedure

The sample consisted of children and adolescents who were referred to a specialized outpatient mental health center in The Netherlands that provides diagnostic and treatment facilities for children and adolescents with a broad range of psychiatric and psychosocial problems. All participants were subjected to an extensive diagnostic procedure according to the longitudinal-expert-all data (LEAD) principle (Spitzer 1983).

Table 2 Demographic and clinical characteristics of the participants in the ASD group (as defined by clinical DSM-IV diagnosis, ADI-R supported DSM-IV diagnosis, and ADI-R supported DSM-5 diagnosis) and the corresponding clinical control group

	Clinical DSM-IV diagnosis ASD group (n = 75)	Clinical control group (n = 57)	ADI-R supported DSM-IV diagnosis ASD group (n = 48)	Clinical control group (n = 49)	ADI-R supported DSM-5 diagnosis ASD group (n = 57)	Clinical control group (n = 46)
Mean age (SD)	11.53 (3.49)	11.75 (3.10)	11.84 (3.68)	11.79 (2.93)	11.36 (3.56)	12.16 (3.05)
Boys (%)	84%	72%	83%	74%	84%	72%
Average IQ (SD)†	102.52 (14.50)	97.19 (14.31)	102.52 (15.46)	98.73 (13.79)	103.41 (15.25)	98.84 (14.72)
Frequency of diagnoses						
Autism spectrum disorder††	75	0	48	0	57	0
Asperger's disorder	14	0	7	0		
Autistic disorder	8	0	7	0		
PDD-NOS	53	0	34	0		
ADHD	31	40	19	35	23	33
Anxiety disorders	9	16	5	12	6	11
Disruptive, impulse-control, and conduct disorders	3	5	2	6	3	4
Specific Learning disorder	12	8	9	8	10	6
Depressive disorders	7	6	6	5	6	6
Relational problems	30	33	17	32	22	26
Communication disorders	4	1	3	1	3	1
Other diagnosis	9	14	5	9	6	10
Diagnosis deferred	1	7	0	6	1	2

ASD Autism Spectrum Disorder, DSM-IV Diagnostic and Statistical Manual of Mental Disorders, fourth edition, ADI-R Autism Diagnostic Interview-Revised, DSM-5 Diagnostic and Statistical Manual of Mental Disorders, fifth edition, PDD-NOS Pervasive developmental disorder-not otherwise specified, ADHD Attention-deficit/hyperactivity disorder. †IQ was only available for 89 participants. ††In DSM-IV defined as Pervasive developmental disorders

This means that diagnoses were made, and revised when necessary, by a multidisciplinary team that included licensed psychiatrists and psychologists, using information from multiple sources (i.e., child, parent, teachers) and multiple assessments (i.e., intake interviews, psychiatric examination, psychological assessment, and observations) during various stages of the clinical process. The CBCL and TRF were routinely completed as part of the intake assessment in order to screen for relevant problem areas, but these instruments were not decisive in the diagnostic process.

Most of the data (66%) were collected by examining existing case records. Cases referred to the outpatient center between 2011 and 2016 were included if the ADI-R had been administered during the diagnostic process, and CBCL and/or TRF data were available. Note that for these cases the ADI-R had already been administered during the diagnostic process because the presence of ASD was hypothesized. To further enlarge the ASD and control groups, additional parents of youth with ASD (who for some reason had not been tested with the ADI-R during the diagnostic process, for example because ASD had already been diagnosed in the past by another mental health institution) or ADHD were actively recruited for the study (between July 2014 and February 2016). Thus, in these cases the ADI-R was completed solely for the purpose of this study, which is why these parents received a small incentive (a voucher) for their participation. The study was approved by the local ethical review committee and participation only occurred after informed consent had been given.

Assessments

As noted in the introduction, the school-aged versions of the *CBCL* and *TRF* of the *ASEBA instrument* (Achenbach and Rescorla 2001) are widely employed standardized scales to assess competences and problems in children and adolescents aged 6 to 18 years. The problems part of these scales contains 113 items referring to behavioral, social, and emotional symptoms, for which parents and teachers are asked to rate the applicability to the child during the past period (parents: 6 months; teachers: 2 months) using a 3-point Likert scale (0 = not true, 1 = somewhat or sometimes true, and 2 = very true or often true). The psychometric properties of the CBCL are well-established in clinical and typically developing youths (Achenbach and Rescorla 2001; Bérubé and Achenbach 2010), and for children and adolescents with ASD (Pandolfi et al. 2012). In the present study, three separate syndrome scales were used as they are considered as particularly relevant for ASD: Withdrawn/depressed (CBCL $\alpha = .76$; TRF $\alpha = .68$), thought problems (CBCL $\alpha = .69$; TRF $\alpha = .68$), and social problems (CBCL $\alpha = .77$; TRF $\alpha = .68$). Two combinations of these existing subscales were also calculated: The ASD profile of Biederman et al. (2010; CBCL $\alpha = .82$; TRF $\alpha = .76$) and the WTP scale of Havdahl et al.

(2016; CBCL $\alpha = .77$; TRF $\alpha = .69$). Finally, two special ASD scales were employed: The ASD scale of Ooi et al. (2011; CBCL $\alpha = .61$; TRF $\alpha = .61$) and the ASD scale of So et al. (2013; CBCL $\alpha = .57$; TRF $\alpha = .57$).

The ADI-R (Rutter et al. 2003; Dutch version by De Jonge and De Bildt 2007) is a comprehensive, semi-structured diagnostic interview that is used to assess the typical features of ASD. The interview consists of 93 items referring to youth's developmental history and current behaviors in three domains: (1) language and communication, (2) reciprocal social interactions, and (3) restricted, repetitive, and stereotyped behaviors and interests. The ADI-R was administered to the parents of the children and adolescents by experienced psychologists who had been trained in conducting this interview. Parents' responses were scored by the interviewer with a score of 0 indicating the absence of a symptom, or a score of 1, 2, or 3 signaling the presence and specifying the severity of a symptom. The psychometric properties of the ADI-R are well-established (Lord et al. 1994). To verify the presence or absence of a DSM-IV diagnosis of ASD, the diagnostic algorithm of the ADI-R consisting of 42 items was applied. In order to confirm a DSM-IV ASD diagnosis, participants had to meet the cut-off for the social domain and either the communication or the repetitive domain; or meet criteria of the social or communication domain and having a required minimal number of symptoms related to the criteria of other domains (Risi et al. 2006). To confirm the DSM-5 diagnosis, the ADI-R was used following the scoring guideline as described by Huerta and colleagues (Huerta et al. 2012), which implies that participants have to score positively on at least one symptom of all three criteria referring to persistent deficits in social communicational and social interaction, and at least one symptom for two out of four criteria regarding restricted repetitive patterns of behavior, interests, or activities.

Data Analysis

In order to explore which of the five screening variants of the ASEBA is best suitable to predict the presence of ASD, a series of logistic regression analyses was conducted. In these analyses, raw scores on the CBCL and TRF scores were the predictors, while ASD diagnosis as obtained in three ways (i.e., clinical DSM-IV diagnosis, ADI-R supported DSM-IV diagnosis, and ADI-R supported DSM-5 diagnosis) was the dependent variable. For each outcome variable, separate analyses were conducted for the 5 screening variants (i.e., separate syndrome scales, the ASD profile, the WTP scale, and the two special ASD scales) and for the 3 informants (parent, teacher, and combined). Because we employed raw ASEBA scores, age and gender were included as covariates in the analyses.

In order to examine and compare the ability of the ASEBA screening variants to correctly classify children with and without ASD, Area Under the Curve (AUC) scores were calculated. The AUC is the area under the Receiver Operating

Characteristic (ROC) curve which plots the true positive rate (sensitivity) against the false positive rate (specificity) for all possible cutoff points. The AUC ranges between 0.5 (chance level) and 1.0 (perfect fit). AUC scores between 0.50 and 0.69 can be considered as poor, scores between 0.70 and 0.79 as fair, scores between 0.80 and 0.89 as good, and scores of 0.90 and higher as excellent (Ferdinand 2008).

For the best ASD screening variants, the most optimal cut-off points were established. These cut-off points were based on a statistical optimal equilibrium between sensitivity and specificity. For each scale and corresponding cut-off point, the sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) were calculated. In order to test the hypothesized added value of combining CBCL and TRF scores (as compared to the use of only one informant), DeLong et al. (1988) test for pairwise comparison of ROC curves was applied.

Results

Predicting the Clinical DSM-IV Diagnosis

CBCL As can be seen in the upper panel of Table 3, the results showed that the social problems and the thought problems syndrome scales on their own did not discriminate between the ASD and clinical control group. The withdrawn/depressed scale, the ASD profile (Biederman et al. 2010) and the WTP scale (Havdahl et al. 2016) did show statistically significant effects, but the OR's ranging from 1.07 to 1.22 were fairly small and the associated AUC scores (0.67–0.68) were relatively low. The results for the ASD scale of Ooi and colleagues (Ooi et al. 2011) and the ASD scale of So and colleagues (So et al. 2013) were more convincing, with OR's being 1.38 and 1.68, respectively and fair AUC scores of 0.74 and 0.78.

TRF The analyses performed on the TRF revealed that neither the separate syndrome scales nor their combinations yielded significant results (see middle panel of Table 3). Again, So et al.'s (2013) and Ooi et al.'s (2011) ASD scales did significantly discriminate between youth with and without ASD, although the OR's of 1.20 and 1.26 were quite modest and the AUC scores of 0.67 and 0.66 should be interpreted as poor.

CBCL & TRF When combining the CBCL and TRF, only the WTP scale (Havdahl et al. 2016) and the ASD scales of Ooi et al. (2011) and So et al. (2013) were found to significantly discriminate between youth with and without ASD (see lower panel of Table 3). When looking at the OR's the two specially developed ASD scales seemed to perform better than the WTP profile, although the AUC scores for all three models were in the fair range.

Predicting the ADI-R Supported DSM-IV Diagnosis

CBCL The analyses for the ADI-R supported DSM-IV diagnosis revealed statistically significant effects for all scales except for thought problems (see upper panel of Table 4). The OR's varied from 1.10 to 1.62 and AUC scores ranged between 0.67 and 0.83. In line with the previously described results, the ASD scales of Ooi et al. (2011) and So et al. (2013) produced the highest OR's and AUC scores well above .80.

TRF The analyses using TRF scores as predictor of ADI-R supported DSM-IV diagnosis involved a smaller number of participants and were therefore subject to power problems. This might explain why both the logistic regression analyses and the ROC analyses yielded no significant effects (see middle panel of Table 4). Thus, none of the proposed TRF scales was able to predict an ADI-R supported DSM-IV diagnosis.

CBCL & TRF Despite the small sample size, the combined CBCL and TRF ASD scales of Ooi et al. (2011) and So et al. (2013) significantly discriminated between children with ASD and the clinical comparison group. The OR's were respectively 1.29 and 1.31 and the accompanying AUC scores 0.77 and 0.78 (see lower panel of Table 4). Other scales did not significantly distinguish between youth with and without ASD.

Predicting the ADI-R Supported DSM-5 Diagnosis

CBCL The logistic regression analyses testing the various CBCL scales as predictors of an ADI-R supported DSM-5 diagnosis of ASD (see upper panel of Table 5) revealed that, with the exception of the thought problems syndrome scale, all scales differentiated between youth with and without ASD. However, note that the OR's were quite divergent with values ranging from 1.11 to 1.75, and an inspection of the AUC scores indicated that especially the ASD scale of So et al. (2013) had the best predictive value with an AUC score of 0.86, closely followed by the ASD scale of Ooi et al. (2011) which had an AUC score of .83.

TRF Only the the ASD scale of Ooi et al. (2011) was able to significantly predict the ADI-R supported DSM-5 diagnosis, with an OR of 1.54 and a fair AUC score of 0.75 (see middle panel of Table 5).

CBCL & TRF The results of the logistic regression analyses using the combined CBCL and TRF scales as predictors again showed that the ASD scales of Ooi et al. (2011) and So et al. (2013) significantly discriminated between youth with and without ASD (see lower panel of Table 5). OR's were 1.44 and 1.57, respectively, and the corresponding AUC scores of 0.81 and 0.86 were good. Individual

Table 3 Potential of various ASEBA screeners for identifying youth with ASD as defined by Clinical DSM-IV Diagnosis: Mean scores of ASD versus clinical control group and results of logistic regression analyses predicting ASD diagnosis and ROC analyses

Operationalization of ASEBA screener	<i>M (SD)</i>		Logistic regression				ROC analyses		
	ASD group	Clinical control group	<i>B</i>	Wald	OR	95% CI for OR	AUC	95% CI for AUC	<i>p</i>
CBCL (<i>N</i> = 114)									
1. Separate syndrome scales:									
Withdrawn/depressed	5.6 (3.7)	3.8 (2.9)	.20**	9.62	1.22	1.08–1.38	.67	.57–.78	.002
Thought problems	5.2 (4.0)	3.9 (3.4)	.09	2.91	1.10	.99–1.22	.65	.55–.75	.007
Social problems	5.1 (3.9)	4.1 (3.4)	.09	2.62	1.09	.98–1.21	.64	.54–.75	.008
2. ASD profile of Biederman et al. (2010)	15.5 (7.7)	11.6 (8.0)	.07*	6.26	1.07	1.01–1.12	.68	.58–.78	.001
3. WTP scale of Havdahl et al. (2016)	10.5 (5.8)	7.7 (5.6)	.09*	6.19	1.09	1.02–1.17	.68	.59–.78	.001
4. ASD scale of Ooi et al. (2011)	4.9 (2.7)	2.8 (2.4)	.33***	16.40	1.38	1.18–1.62	.74	.65–.83	<.001
5. ASD scale of So et al. (2013)	6.5 (3.0)	3.8 (2.6)	.37***	20.29	1.68	1.23–1.72	.78	.70–.87	<.001
TRF (<i>N</i> = 79)									
1. Separate syndrome scales:									
Withdrawn/depressed	4.3 (3.2)	3.6 (2.4)	.13	2.02	1.14	.95–1.36	.63	.50–.76	.061
Thought problems	2.0 (2.6)	1.3 (2.2)	.16	1.84	1.17	.93–1.47	.65	.52–.78	.035
Social problems	3.3 (3.1)	4.1 (3.2)	–.05	.42	.95	.82–1.11	.63	.50–.76	.068
2. ASD profile of Biederman et al. (2010)	9.5 (6.2)	8.9 (5.9)	.04	.73	1.04	.95–1.13	.62	.49–.76	.078
3. WTP scale of Havdahl et al. (2016)	6.3 (4.3)	4.9 (3.6)	.12	2.92	1.12	.98–1.28	.66	.53–.79	.025
4. ASD scale of Ooi et al. (2011)	3.3 (2.7)	2.3 (2.6)	.23*	4.35	1.26	1.01–1.56	.66	.53–.79	.016
5. ASD scale of So et al. (2013)	4.9 (3.3)	3.7 (2.2)	.18*	3.98	1.20	1.00–1.44	.67	.55–.79	.024
CBCL & TRF (<i>N</i> = 66)									
1. Separate syndrome scales:									
Withdrawn/depressed	9.7 (5.8)	7.7 (4.1)	.10	3.32	1.11	.99–1.23	.68	.54–.81	.019
Thought problems	7.3 (5.9)	5.3 (3.8)	.13	3.70	1.14	1.00–1.30	.72	.60–.85	.003
Social problems	8.4 (6.4)	9.4 (5.2)	–.01	.05	.99	.91–1.08	.64	.50–.78	.066
2. ASD profile of Biederman et al. (2010)	24.8 (12.2)	22.6 (10.2)	.04	2.00	1.04	.99–1.09	.68	.55–.81	.017
3. WTP scale of Havdahl et al. (2016)	16.7 (8.9)	13.1 (6.6)	.09**	5.00	1.09	1.01–1.18	.74	.62–.87	.001
4. ASD scale of Ooi et al. (2011)	8.3 (4.4)	5.7 (3.2)	.23**	8.44	1.26	1.08–1.48	.77	.65–.89	<.001
5. ASD scale of So et al. (2013)	11.4 (5.1)	7.9 (3.3)	.27**	10.34	1.31	1.11–1.55	.80	.70–.90	<.001

ASEBA Achenbach Scales of Empirically Based Assessment, ASD Autism Spectrum Disorder, DSM-IV Diagnostic and Statistical Manual of Mental Disorders, fourth edition, ROC Receiver Operating Characteristics, OR Odds Ratio, CI Confidence Interval, AUC Area Under the Curve, CBCL Child Behavior Checklist, TRF Teacher Report Form, WTP Withdrawn/depressed and Thought problems. * $p < .05$, ** $p < .01$, *** $p < .001$

syndrome scales as well as their combinations did not show significant effects.

Additional Examination of the Best Screening Scales

Overall, the results demonstrated that the ASD scales of Ooi et al. (2011) and So et al. (2013) were best in discriminating between ASD and clinical control youth, and thus had most predictive power for the diagnosis of this disorder (as established in various ways). In order to further examine the screening potential of these two scales, we determined the most optimal cut-off points and explored their sensitivity, specificity, positive predictive value, and negative predictive value. The optimal cut-off points and the relevant statistics are

presented in Table 6. As can be seen, results were highly comparable for both ASD scales. That is, their predictive validity was most optimal in case the CBCL data were used. For the ASD scale of Ooi et al. (2011), the CBCL was capable of correctly identifying between 73 to 83% of the children with ASD (sensitivity) and between 69 to 72% of the children without the disorder (specificity). In addition, 71 to 75% of the children and adolescents meeting the cut-off score on the CBCL was indeed diagnosed with ASD (PPV) and 65 to 79% of the children and adolescents who did not meet the cut-off indeed did not have the disorder (NPV). For the ASD scale of So et al. (2013) similar figures were found. Thus, when employing the CBCL, between 72 to 76% of the children and adolescents were correctly identified with ASD

Table 4 Potential of various ASEBA screeners for identifying youth with ASD as defined by ADI-R supported DSM-IV diagnosis: Mean scores of ASD versus clinical control group and results of logistic regression analyses predicting ASD diagnosis and ROC analyses

Operationalization of ASEBA screener	<i>M (SD)</i>		Logistic regression				ROC analyses		
	ASD group	Clinical control group	<i>B</i>	Wald	OR	95% CI for OR	AUC	95% CI for AUC	<i>p</i>
<i>CBCL (N = 86)</i>									
1. Separate syndrome scales:									
Withdrawn/depressed	6.33 (3.94)	3.53 (2.96)	.27***	12.65	1.31	1.13–1.52	.73	.62–.84	<.001
Thought problems	4.78 (3.50)	3.67 (3.09)	.10	2.24	1.11	.97–1.27	.64	.52–.75	.029
Social problems	5.38 (4.23)	3.43 (2.89)	.16*	6.34	1.18	1.04–1.33	.67	.56–.79	.006
2. ASD profile of Biederman et al. (2010)	15.98 (7.94)	10.38 (7.27)	.10**	9.17	1.10	1.03–1.17	.72	.61–.83	<.001
3. WTP scale of Havdahl et al. (2016)	10.80 (5.81)	7.16 (5.42)	.12**	7.41	1.13	1.03–1.22	.71	.60–.82	.001
4. ASD scale of Ooi et al. (2011)	5.42 (2.71)	2.58 (2.44)	.42***	18.40	1.52	1.25–1.83	.80	.70–.89	<.001
5. ASD scale of So et al. (2013)	6.82 (3.07)	3.38 (2.33)	.48***	21.04	1.62	1.32–2.00	.83	.74–.92	<.001
<i>TRF (N = 54)</i>									
1. Separate syndrome scales:									
Withdrawn/depressed	4.66 (3.20)	3.76 (2.56)	.12	1.32	1.12	.92–1.37	.63	.48–.77	.128
Thought problems	1.51 (1.70)	1.18 (1.91)	.12	.47	1.12	.81–1.57	.63	.48–.78	.115
Social problems	3.11 (3.31)	3.86 (2.82)	-.04	.20	.96	.80–1.15	.61	.46–.76	.191
2. ASD profile of Biederman et al. (2010)	8.92 (5.62)	8.62 (5.30)	.02	.18	1.02	.92–1.14	.64	.49–.79	.092
3. WTP scale of Havdahl et al. (2016)	5.95 (3.57)	4.92 (3.30)	.10	1.18	1.10	.93–1.31	.65	.51–.80	.060
4. ASD scale of Ooi et al. (2011)	3.38 (2.87)	2.31 (2.22)	.22	3.08	1.25	.97–1.60	.65	.51–.80	.060
5. ASD scale of So et al. (2013)	4.66 (3.29)	3.73 (2.19)	.15	1.81	1.16	.94–1.44	.65	.50–.79	.076
<i>CBCL & TRF (N = 46)</i>									
1. Separate syndrome scales:									
Withdrawn/depressed	11.19 (6.11)	7.97 (4.41)	.11	3.42	1.12	.99–1.26	.69	.54–.84	.030
Thought problems	6.12 (3.89)	5.45 (3.88)	.08	.97	1.09	.91–1.29	.66	.50–.82	.069
Social problems	8.19 (6.88)	8.45 (4.38)	.02	.09	1.02	.92–1.12	.64	.48–.81	.101
2. ASD profile of Biederman et al. (2010)	24.73 (11.86)	22.00 (10.47)	.04	1.28	1.04	.98–1.10	.67	.51–.83	.048
3. WTP scale of Havdahl et al. (2016)	16.94 (8.23)	13.53 (7.06)	.07	2.40	1.08	.98–1.18	.73	.58–.88	.009
4. ASD scale of Ooi et al. (2011)	8.88 (4.49)	5.78 (3.42)	.25*	6.70	1.29	1.06–1.56	.77	.63–.91	.002
5. ASD scale of So et al. (2013)	11.47 (5.07)	7.77 (3.35)	.27**	7.20	1.31	1.08–1.59	.78	.65–.91	.001

ASEBA Achenbach Scales of Empirically Based Assessment, ASD Autism Spectrum Disorder, DSM-IV Diagnostic and Statistical Manual of Mental Disorders, fourth edition, ROC Receiver Operating Characteristics, OR Odds Ratio, CI Confidence Interval, AUC Area Under the Curve, CBCL Child Behavior Checklist, TRF Teacher Report Form, WTP Withdrawn/depressed and Thought problems. * $p < .05$, ** $p < .01$, *** $p < .001$

(sensitivity) and 65 to 71% of them were identified as not having the disorder (specificity). Further, 71 to 76% of the youth meeting the cut-off score were indeed diagnosed with ASD (PPV), while 64 to 75% of the youth who did not meet the cut-off score indeed were not diagnosed with the disorder (NPV).

On first sight, the Ooi et al. (2011) and So et al. (2013) scales based on the TRF or the combined CBCL and TRF data appeared less useful than the ASD scales based on the CBCL. As can be seen in Table 6, in particular the specificity and NPV decreased when the TRF or combined CBCL/TRF data were employed. However, it should be noted that DeLong et al. (1988) test for pairwise comparison of ROC curves could not substantiate this apparent superiority of the CBCL-based

over the TRF-based ASD scales of Ooi et al. (2011) [$Z = .51, p = .61$] and So et al. (2013) [$Z = 1.07, p = .29$]. In a similar vein, no significant differences emerged between the CBCL-based ASD scales and the combined CBCL and TRF-based ASD scales of Ooi et al. (2011) [$Z = .64, p = .54$] and So et al. (2013) [$Z = .27, p = .79$] respectively.

Discussion

A screening possibility for ASD within the widely applied ASEBA scales CBCL and TRF could be of great value for daily clinical practice as this might facilitate the diagnostic process without additional costs, time, and effort. In the

Table 5 Potential of various ASEBA screeners for identifying youth with ASD as defined by ADI-R supported DSM-5 diagnosis: Mean scores of ASD versus clinical control group and results of logistic regression analyses predicting ASD diagnosis and ROC analyses

Operationalization of ASEBA screener	<i>M (SD)</i>		Logistic regression				ROC analyses		
	ASD group	Clinical control group	<i>B</i>	Wald	OR	95% CI for OR	AUC	95% CI for AUC	<i>p</i>
<i>CBCL (N = 88)</i>									
1. Separate syndrome scales:									
Withdrawn/depressed	5.78 (3.73)	3.60 (2.92)	.26**	11.24	1.29	1.11–1.50	.73	.63–.84	<.001
Thought problems	4.80 (3.43)	3.31 (2.82)	.14	3.6	1.16	1.00–1.34	.68	.57–.79	.004
Social problems	5.57 (4.12)	3.58 (3.35)	.14*	4.86	1.15	1.02–1.29	.69	.58–.80	.002
2. ASD profile of Biederman et al. (2010)	15.78 (7.42)	10.23 (7.12)	.10**	9.33	1.11	1.04–1.19	.73	.63–.84	<.001
3. WTP scale of Havdahl et al. (2016)	10.43 (5.41)	6.87 (5.02)	.14**	8.38	1.15	1.05–1.26	.73	.63–.84	<.001
4. ASD scale of Ooi et al. (2011)	5.37 (2.46)	2.57 (2.41)	.44***	19.62	1.56	1.28–1.90	.83	.74–.91	<.001
5. ASD scale of So et al. (2013)	6.74 (2.81)	3.23 (2.24)	.56***	22.11	1.75	1.39–2.21	.86	.79–.94	<.001
<i>TRF (N = 60)</i>									
1. Separate syndrome scales:									
Withdrawn/depressed	4.28 (3.13)	3.61 (2.33)	.11	1.03	1.11	.91–1.37	.60	.44–.76	.247
Thought problems	1.85 (2.39)	1.21 (2.25)	.14	.86	1.15	.86–1.55	.63	.47–.78	.133
Social problems	3.43 (3.36)	3.83 (3.26)	-.02	.07	.98	.82–1.17	.60	.44–.65	.254
2. ASD profile of Biederman et al. (2010)	9.27 (5.70)	8.44 (5.71)	.04	.41	1.04	.93–1.15	.60	.44–.76	.228
3. WTP scale of Havdahl et al. (2016)	5.95 (3.65)	4.80 (3.32)	.11	1.55	1.12	.94–1.33	.63	.47–.79	.117
4. ASD scale of Ooi et al. (2011)	3.33 (2.59)	3.83 (3.26)	.43*	5.88	1.54	1.09–2.17	.75	.61–.90	.002
5. ASD scale of So et al. (2013)	4.60 (2.97)	3.09 (1.88)	.25	3.82	1.29	1.00–1.66	.68	.54–.81	.035
<i>CBCL & TRF (N = 48)</i>									
1. Separate syndrome scales:									
Withdrawn/depressed	10.03 (5.84)	7.97 (4.43)	.10	2.60	1.11	.98–1.25	.68	.52–.84	.049
Thought problems	6.40 (4.67)	4.77 (3.77)	.12	1.80	1.13	.95–1.35	.68	.53–.84	.042
Social problems	8.83 (6.90)	9.00 (5.49)	.00	.00	1.00	.91–1.11	.62	.45–.79	.186
2. ASD profile of Biederman et al. (2010)	24.44 (11.17)	21.94 (10.79)	.03	1.03	1.03	.97–1.10	.68	.52–.84	.049
3. WTP scale of Havdahl et al. (2016)	16.11 (7.54)	12.88 (6.94)	.08	2.73	1.09	.99–1.20	.72	.56–.87	.018
4. ASD scale of Ooi et al. (2011)	8.67 (3.86)	4.91 (2.91)	.37**	8.87	1.44	1.13–1.83	.81	.67–.95	.001
5. ASD scale of So et al. (2013)	11.31 (4.26)	6.90 (3.08)	.45***	9.66	1.57	1.18–2.08	.86	.75–.97	<.001

ASEBA Achenbach Scales of Empirically Based Assessment, ASD Autism Spectrum Disorder, DSM-5 Diagnostic and Statistical Manual of Mental Disorders, fifth edition, ROC Receiver Operating Characteristics, OR Odds Ratio, CI Confidence Interval, AUC Area Under the Curve, CBCL Child Behavior Checklist, TRF Teacher Report Form, WTP Withdrawn/depressed and Thought problems. * $p < .05$, ** $p < .01$, *** $p < .001$

present study, five different screening variants were compared using CBCL, TRF, and combined CBCL and TRF data within a clinically referred sample.

When using the CBCL as a screen for ASD, the results demonstrated that the special ASD scales of Ooi et al. (2011) and So et al. (2013) had the best potential to discriminate children with ASD (as diagnosed in various ways) from children without ASD, with odds ratios and AUC values being clearly superior to those found for screens based on individual or combinations of CBCL syndrome scales. A similar conclusion was true for the TRF: Again the special ASD scales of Ooi et al. (2011) and So et al. (2013) had the best screening potential, although it should be noted that odds ratios and AUC values were somewhat lower than those obtained for

the CBCL. Note that this pattern of findings was robust and not dependent on the way ASD had been defined (i.e., clinical DSM-IV diagnosis or ADI-R supported DSM-IV or DSM-5 diagnosis). Altogether, these findings confirm our first hypothesis that the special ASD scales as construed by Ooi et al. (2011) and So et al. (2013) are better in screening for ASD than the existing syndrome scales of the ASEBA instrument.

In passing, it should be mentioned that our data suggest that parents (CBCL) are better informants when diagnosing youth with ASD than teachers (TRF). In a way, this makes sense as parents can observe their child in wide range of situations and often – via communication with the teacher – are also well aware of the child's behavior and functioning at school. In

Table 6 Sensitivity, specificity, PPV, and NPV (rounded off) percentages for the optimal cut-off points found for the ASEBA screening scales as proposed by Ooi et al. (2011) and So et al. (2013) for identifying ASD (as defined by clinical DSM-IV diagnosis, ADI-R supported DSM-IV diagnosis, and ADI-R supported DSM-5 diagnosis)

	Cut-off scores	Clinical DSM-IV diagnosis				ADI-R supported DSM-IV diagnosis				ADI-R supported DSM-5 diagnosis			
		Sens.	Spec.	PPV	NPV	Sens.	Spec.	PPV	NPV	Sens.	Spec.	PPV	NPV
Ooi et al. (2011)													
CBCL	4	73	72	79	65	80	69	71	79	83	71	76	78
TRF	2	71	46	75	41	71	43	68	47	75	73	89	50
CBCL & TRF	7	69	63	80	47	72	60	74	57	73	69	86	50
So et al. (2013)													
CBCL	5	72	65	73	64	76	71	71	75	76	70	76	70
TRF	4	60	56	76	38	57	60	71	44	59	71	85	40
CBCL & TRF	10	61	65	80	42	63	63	74	50	63	73	87	42

ASEBA Achenbach Scales of Empirically Based Assessment, DSM-IV Diagnostic and Statistical Manual of Mental Disorders, fourth edition, DSM-5 Diagnostic and Statistical Manual of Mental Disorders, fifth edition, Sens. sensitivity, Spec. specificity, PPV positive predictive value, NPV negative predicted value, CBCL Child Behavior Checklist, TRF Teacher Report Form

contrast, the teacher’s information is mainly based on what can be observed at school; he/she has little information about the child’s behaviors in other situations. Meanwhile, it may well be that the methodology of the present study was more ‘in favor’ of parents than of teachers, because the former were obviously more involved in the diagnostic process (e.g., administration of the ADI-R) than the latter.

When using the combined CBCL and TRF data, the special ASD scales again produced the best outcomes. However, no evidence was obtained for our second hypothesis that the employment of combined CBCL and TRF data would enhance the screening potential as compared to using ASEBA scale data of only one informant. This is in contrast with So et al. (2013) who found that the predictive power for detecting ASD significantly increased when taking the TRF into account. Note, however, that even in the So et al. (2013) study the additional value of including the TRF in the screening procedure was also quite limited. Depending on the comparison group used, the combined CBCL- and TRF-based ASD scale yielded 1–8% increase in the accuracy of predicting ASD as compared to the ASD scale based on the CBCL or TRF alone.

It should be mentioned, that the ASD scales of Ooi et al. (2011) and So et al. (2013) display considerable overlap in terms of item content. The six common items of the CBCL/TRF that are included in both versions appear to cover delayed development/communication (i.e., “Acts too young for his/her age” and “Speech problem”), social problems (i.e., “Would rather be alone than with others” and “Withdrawn, doesn’t get involved with others”), and awkward/stereotyped behaviors (i.e., “Repeats certain acts over and over; compulsions” and “Strange behavior”), which nicely capture the key symptoms of ASD (American Psychiatric Association 2000, 2013).

Not surprisingly, the screening potential of both scales appeared to be highly comparable.

The predictive value of the ASD scales of Ooi et al. (2011) and So et al. (2013) is comparable to that of the standard syndrome and DSM-oriented scales of the Achenbach scales for detecting other types of child and adolescent psychopathology such as anxiety disorders, depression, ADHD, and disruptive behavior disorders (Ebesutani et al. 2010; Ferdinand 2008). However, in absolute terms, the screening potential of these ASD scales appears limited. That is, the sensitivity, specificity, PPV, and NPV percentages were clearly inferior to those obtained for specific screening instruments such as the SCQ, SRS, and CCC that can be employed for detecting ASD in youth with IQ > 70 (Auyeung et al. 2008; Charman et al. 2007). However, given their quite good sensitivity and PPV percentages, the ASEBA-based ASD scales as proposed by Ooi et al. (2011) and So et al. (2013) – especially when parents are used as informant, seem to be useful as a first step in the diagnostic process. Subsequently, the more specific instruments could be administered, before launching more elaborate diagnostic tests such as the ADI-R and the ADOS-2.

The present study suffers from a number of limitations. First, the sample size was relatively small and this was especially true for the teacher report data, which may have resulted in limited power for testing the predictive value for ASD by means of the Achenbach scales. Second, the present study made use of a convenience sample of children and adolescents who were referred to a regular outpatient treatment facility. This implied that most cases were “diagnostic puzzles” for which the ASD picture was not particularly clear. In other words, most youth who were ultimately diagnosed with ASD were not showing symptoms on the extreme end of the spectrum, whereas at least part of the youth in the clinical control group at least showed some ASD-like characteristics.

In addition, most youth in the control group had a primary diagnosis of ADHD, a developmental disorder showing considerable comorbidity and overlap in symptoms and problems with ASD (Ronald et al. 2014; Taurines et al. 2012). In other words, the composition of this sample probably made it more difficult to distinguish between youth with and without ASD, and as such the screening potential of the ASEBA scales might be even better than the current findings suggest. Third and finally, the reliability of the special ASD scales of Ooi et al. (2011) and So et al. (2013) appeared to be quite modest. That is, previous studies documented Cronbach's alphas in the .56 to .78 range (Ooi et al. 2011; So et al. 2013), whereas in the present investigation internal consistency coefficients between .57 and .61 were found. This may well have to do with the fact that ASD symptoms are quite heterogeneous (Masi et al. 2017): they include social-communicative difficulties as well as restrictive and repetitive patterns of behavior and individual children and adolescents with this condition can be quite different in terms of symptomatology. Moreover, as noted above, the present sample included few severe cases of ASD, which may also have undermined the emergence of more substantial correlations among items, which is a prerequisite for finding a high internal consistency coefficient for the special ASD scales.

In spite of these shortcomings, the present study replicated that the ASEBA-based scales as proposed by Ooi et al. (2011) and So et al. (2013) – especially when completed by the parent – are valuable screens for ASD in a clinically referred sample. These special ASD scales turned out to be better predictors of ASD than the existing syndrome scales and their combinations, and this was independent of the way the diagnosis was established (i.e., clinical versus ADI-R supported diagnosis, DSM-IV versus DSM-5). Clearly more research with large, representative, and well-described clinical and non-clinical samples is needed to definitively evaluate the utility of the CBCL and TRF as screening instruments for ASD. More precisely, based on the results of the present study, it is clear that future research should concentrate on the two special ASD scales of Ooi et al. (2011) and So et al. (2013). In addition, a major point of interest for future research is determining adequate cut-off scores. There is an inversely proportional relation between the sensitivity and the specificity. Increasing the sensitivity and decreasing false negatives is always at the expense of the specificity and false positives. The optimal cut-off score is highly dependent on the purpose and context of the instrument. Within a referred sample, high sensitivity would be preferred when the ASEBA is used as an initial screen. However, such a strategy seems less useful when screening the general population as it may result in the detection (and referral) of many false positives.

Although additional research is needed, the results have clear clinical implications as they underscore the potential of an ASD subscale included in the Achenbach scales. This ASD

subscale could be a good, easy-to-administer initial screen for detecting this type of psychopathology. Because of the wide use of the ASEBA instrument, no extra costs have to be made. The results also indicate that this is only an initial screen and that further diagnostic evaluation (i.e., ADI-R, ADOS-2) is needed to definitively establish the diagnosis. Meeting the cut-off score should alert clinicians of the possibility of ASD and encourage them to include ASD in their differential diagnostic consideration and their decision making with regard to further diagnostic procedures and/or referral strategy.

Compliance with Ethical Standards

Conflict of Interest The authors state that there is no conflict of interest.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Achenbach, T. M. (1991). *Manual for the child behavior checklist/4–18 and 1991 profile*. Burlington: University of Vermont Department of Psychiatry.
- Achenbach, T. M., & Rescorla, L. (2000). *Manual for the ASEBA preschool forms and profiles: An integrated system of multi-informant assessment*. Burlington: University of Vermont, Department of Psychiatry.
- Achenbach, T. M., & Rescorla, L. A. (2001). *Manual for the ASEBA school-age forms & profiles*. Burlington, VT: University of Vermont, Research Center for Children, Youth, and Families.
- American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders (4th edition, text revision)*. Washington, DC: APA.
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (5th edition)*. Washington, DC: APA.
- Auyeung, B., Baron-Cohen, S., Wheelwright, S., & Allison, C. (2008). The autism-spectrum quotient: Children's version (AQ-child). *Journal of Autism and Developmental Disorders*, 38(7), 1230–1240.
- Baron-Cohen, S., Hoekstra, R. A., Knickmeyer, R., & Wheelwright, S. (2006). The autism-spectrum quotient (AQ)-adolescent version. *Journal of Autism and Developmental Disorders*, 36(3), 343.
- Bérubé, R. L., & Achenbach, T. M. (2010). *Bibliography of published studies using the Achenbach system of empirically based assessment: 2006 edition*. Burlington: University of Vermont, Research Center for Children, Youth, & Families.
- Berument, S. K., Rutter, M., Lord, C., Pickles, A., & Bailey, A. (1999). Autism screening questionnaire: Diagnostic validity. *British Journal of Psychiatry*, 175, 444–451.
- Biederman, J., Petty, C. R., Fried, R., Wozniak, J., Micco, J. A., Henin, A., Doyle, R., Joshi, G., Galdo, M., Kotarski, M., Caruso, J., Yorks, D., & Faraone, S. V. (2010). Child behavior checklist clinical scales discriminate referred youth with autism spectrum disorder: A preliminary study. *Journal of Developmental and Behavioral Pediatrics*, 31(6), 485–490.

- Bishop, D. V. M. (1998). Development of the Children's communication checklist (CCC): A method for assessing qualitative aspects of communicative impairment in children. *Journal of Child Psychology and Psychiatry*, 39, 879–891.
- Bölte, S., Dickhut, H., & Poustka, F. (1999). Patterns of parent-reported problems indicative in autism. *Psychopathology*, 32(2), 93–97.
- Charman, T., Baird, G., Simonoff, E., Loucas, T., Chandler, S., Meldrum, D., & Pickles, A. (2007). Efficacy of three screening instruments in the identification of autistic-spectrum disorders. *The British Journal of Psychiatry*, 191(6), 554–559.
- Constantino, J. N., & Gruber, C. P. (2005). *Social Responsiveness Scale (SRS)*. Torrance: Western Psychological Services.
- Daniels, A. M., & Mandell, D. S. (2014). Explaining differences in age at autism spectrum disorder diagnosis: A critical review. *Autism*, 18(5), 583–597.
- De Jonge, M. V., & De Bildt, A. (2007). *Autisme diagnostisch interview-revised*. Amsterdam: Hogrefe.
- De Los Reyes, A., & Kazdin, A. E. (2005). Informant discrepancies in the assessment of childhood psychopathology: A critical review, theoretical framework, and recommendations for further study. *Psychological Bulletin*, 131(4), 483.
- DeLong, E. R., DeLong, D. M., & Clarke-Pearson, D. L. (1988). Comparing the areas under two or more correlated receiver operating characteristic curves: A nonparametric approach. *Biometrics*, 44, 837–845.
- Duarte, C. S., Bordin, I. A., de Oliveira, A., & Bird, H. (2003). The CBCL and the identification of children with autism and related conditions in Brazil: Pilot findings. *Journal of Autism and Developmental Disorders*, 33(6), 703–707.
- Ebesutani, C., Bernstein, A., Nakamura, B. J., Chorpita, B. F., Higa-McMillan, C. K., Weisz, J. R., & Research Network on Youth Mental Health. (2010). Concurrent validity of the child behavior checklist DSM-oriented scales: Correspondence with DSM diagnoses and comparison to syndrome scales. *Journal of Psychopathology and Behavioral Assessment*, 32(3), 373–384.
- Falkmer, T., Anderson, K., Falkmer, M., & Horlin, C. (2013). Diagnostic procedures in autism spectrum disorders: A systematic literature review. *European Child and Adolescent Psychiatry*, 22(6), 329–340.
- Ferdinand, R. F. (2008). Validity of the CBCL/YSR DSM-IV scales anxiety problems and affective problems. *Journal of Anxiety Disorders*, 22, 126–134.
- Havdahl, K. A., Von Tetzchner, S., Huerta, M., Lord, C., & Bishop, S. L. (2016). Utility of the child behavior checklist as a screener for autism spectrum disorder. *Autism Research*, 9(1), 33–42.
- Hoffmann, W., Weber, L., König, U., Becker, K., & Kamp-Becker, I. (2016). The role of the CBCL in the assessment of autism spectrum disorders: An evaluation of symptom profiles and screening characteristics. *Research in Autism Spectrum Disorders*, 27, 44–53.
- Howlin, P., & Asgharian, A. (1999). The diagnosis of autism and Asperger syndrome: Findings from a survey of 770 families. *Developmental Medicine and Child Neurology*, 41, 834–836.
- Huerta, M., Bishop, S. L., Duncan, A., Hus, V., & Lord, C. (2012). Application of DSM-5 criteria for autism spectrum disorder to three samples of children with DSM-IV diagnoses of pervasive developmental disorders. *American Journal of Psychiatry*, 169, 1056–1064.
- Kim, Y. S., Leventhal, B. L., Koh, Y. J., Fombonne, E. J., Laska, E., Lim, E. C., Cheon, K., Kim, S., Kim, Y. K., Lee, H., Song, D., & Grinker, R. R. (2011). Prevalence of autism spectrum disorders in a total population sample. *American Journal of Psychiatry*, 168, 904–912.
- Lord, C. E. (2010). Autism: From research to practice. *American Psychologist*, 65(8), 815.
- Lord, C., Rutter, M., & Le Couteur, A. (1994). Autism diagnostic interview-revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *Journal of Autism and Developmental Disorders*, 24(5), 659–685.
- Lord, C., Rutter, M., DiLavore, P., Risi, S., Gotham, K., & Bishop, S. (2012). *Autism diagnostic observation schedule—2nd edition (ADOS-2)*. Los Angeles: Western Psychological Corporation.
- Masi, A., DeMayo, M. M., Glozier, N., & Guastella, A. J. (2017). An overview of autism spectrum disorder, heterogeneity and treatment options. *Neuroscience Bulletin*, 33(2), 183–193.
- Muratori, F., Narzisi, A., Tancredi, R., Cosenza, A., Calugi, S., Saviozzi, I., Santocchi, E., & Calderoni, S. (2011). The CBCL 1.5–5 and the identification of preschoolers with autism in Italy. *Epidemiology and Psychiatric Sciences*, 20(4), 329–338.
- Myers, C. L., Gross, A. D., & McReynolds, B. M. (2014). Broadband behavior rating scales as screeners for autism? *Journal of Autism and Developmental Disorders*, 44(6), 1403–1413.
- Ooi, Y. P., Rescorla, L., Ang, R. P., Woo, B., & Fung, D. S. (2011). Identification of autism spectrum disorders using the child behavior checklist in Singapore. *Journal of Autism and Developmental Disorders*, 41(9), 1147–1156.
- Pandolfi, V., Magyar, C. I., & Dill, C. A. (2012). An initial psychometric evaluation of the CBCL 6–18 in a sample of youth with autism spectrum disorders. *Research in Autism Spectrum Disorders*, 6, 96–108.
- Risi, S., Lord, C., Gotham, K., Corsello, C., Chrysler, C., Szatmari, P., Cook, E. H., Leventhal, B. L., & Pickles, A. (2006). Combining information from multiple sources in the diagnosis of autism spectrum disorders. *Journal of the American Academy of Child and Adolescent Psychiatry*, 45, 1094–1103.
- Ronald, A., Larsson, H., Anckarsäter, H., & Lichtenstein, P. (2014). Symptoms of autism and ADHD: A Swedish twin study examining their overlap. *Journal of Abnormal Psychology*, 123, 440.
- Rutter, M., Le Couteur, A., & Lord, C. (2003). *ADI-R: The autism diagnostic interview-revised*. Los Angeles: Western Psychological Services.
- Sikora, D. M., Hall, T. A., Hartley, S. L., Gerrard-Morris, A. E., & Cagle, S. (2008). Does parent report of behavior differ across ADOS-G classifications: Analysis of scores from the CBCL and GARS. *Journal of Autism and Developmental Disorders*, 38(3), 440–448.
- 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. *Journal of the American Academy of Child and Adolescent Psychiatry*, 47(8), 921–929.
- So, P., Greaves-Lord, K., Van der Ende, J., Verhulst, F. C., Rescorla, L., & De Nijs, P. F. (2013). Using the child behavior checklist and the Teacher's report form for identification of children with autism spectrum disorders. *Autism*, 17(5), 595–607.
- Spitzer, R. L. (1983). Psychiatric diagnosis: Are clinicians still necessary? *Comprehensive Psychiatry*, 24, 399–411.
- Taurines, R., Schwenck, C., Westerwald, E., Sachse, M., Siniatchkin, M., & Freitag, C. (2012). ADHD and autism: Differential diagnosis or overlapping traits? A selective review. *ADHD Attention Deficit and Hyperactivity Disorders*, 4, 115–139.