

Factors influencing scores on the social responsiveness scale

Vanessa Hus,¹ Somer Bishop,² Katherine Gotham,¹ Marisela Huerta,³ and Catherine Lord³

¹Department of Psychology, University of Michigan; ²Cincinnati Children's Hospital Medical Center; ³Department of Psychiatry, Weill-Cornell Medical School

Background: The Social Responsiveness Scale (SRS) is a parent-completed screening questionnaire often used to measure autism spectrum disorders (ASD) severity. Although child characteristics are known to influence scores from other ASD-symptom measures, as well as parent-questionnaires more broadly, there has been limited consideration of how non-ASD-specific factors may affect interpretation of SRS scores. Previous studies have explored effects of behavior problems on SRS specificity, but have not addressed influences on the use of the SRS as a quantitative measure of ASD-symptoms. **Method:** Raw scores (SRS-Raw) from parent-completed SRS were analyzed for 2,368 probands with ASD and 1,913 unaffected siblings. Regression analyses were used to assess associations between SRS scores and demographic, language, cognitive, and behavior measures. **Results:** For probands, higher SRS-Raw were associated with greater non-ASD behavior problems, higher age, and more impaired language and cognitive skills, as well as scores from other parent report measures of social development and ASD-symptoms. For unaffected siblings, having more behavior problems predicted higher SRS-Raw; male gender, younger age, and poorer adaptive social and expressive communication skills also showed small, but significant effects. **Conclusions:** When using the SRS as a quantitative phenotype measure, the influence of behavior problems, age, and expressive language or cognitive level on scores must be considered. If effects of non-ASD-specific factors are not addressed, SRS scores are more appropriately interpreted as indicating general levels of impairment, than as severity of ASD-specific symptoms or social impairment. Additional research is needed to consider how these factors influence the SRS' sensitivity and specificity in large, clinical samples including individuals with disorders other than ASD. **Keywords:** Social Responsiveness Scale, autism spectrum disorder, behavior problems, age, language level.

Autism spectrum disorders (ASD) are characterized by a range of symptoms, which are heterogeneous in nature and severity. A single measure that captured ASD-symptom severity would be useful as a quantitative phenotype in genetic and neurobiological studies. However, in addition to heterogeneity across individuals, measurement of ASD severity is complicated by common co-occurrence of non-ASD-specific conditions (e.g., intellectual disability) and behaviors, such as difficulties with attention or hyperactivity, as well as age-related variation in symptom presentation. Research has demonstrated that raw totals from many ASD diagnostic and screening measures are influenced by non-ASD-specific child characteristics, such as age and language level (e.g., Corsello et al., 2007; Gotham, Pickles, & Lord, 2009; Mayes & Calhoun, 2010). For example, age and language explained 22% of variance in scores from the Autism Diagnostic Interview-Revised (ADI-R; Rutter, Le Couteur, & Lord, 2003), a diagnostic parent-interview often used as a measure of ASD severity (Hus & Lord, 2012).

Recent discussions regarding developmental screening and assessments have called for better

understanding of factors influencing parent report (Aylward, 2009; Warren et al., 2012). Often, parent-factors, such as education level and frame of reference, are acknowledged as limitations and weighed against the relative benefit of efficiency and cost-effectiveness of questionnaires compared to interview or observational measures requiring more time and highly trained clinicians. For example, the Social Responsiveness Scale (SRS; Constantino & Gruber, 2005), a parent-completed questionnaire, which was originally proposed as a continuously distributed, quantitative measure of autism-related severity in the general population (Constantino, Przybeck, Friesen, & Todd, 2000; Constantino et al., 2003a), is commonly used as an estimate of ASD severity in genetic and neurobiological studies.

Although the SRS is frequently referred to as a measure of 'social impairment,' many SRS items describe other core features of ASD, including communication deficits and repetitive behaviors (Constantino et al., 2000), as well as symptoms not exclusively related to ASD diagnostic criteria (Grzadzinski et al., 2011). Informants complete all 65 SRS items, irrespective of the child's age or language level. Without explicit instructions, it remains unclear how parents rate items that are not

Conflict of interest statement: CL receives royalties for the ADI-R and ADOS; profits from this study were donated to charity.

applicable to their child (e.g., items assessing conversation for a nonverbal child). Considering that scores from the ADI-R are affected by child characteristics, despite having subsets of items for children of different ages and language abilities and being administered and scored by a trained clinician, it seems likely that scores on the parent-rated SRS would be similarly influenced. However, in spite of their implications for interpretability of scores, particularly when being used as indicators of ASD-specific severity, studies have not systematically examined how non-ASD-specific child characteristics affect the use of SRS scores as a quantitative measure. The goal of this study was to provide a better understanding of how factors that affect other ASD-symptom measures influence interpretation of SRS scores.

Underscoring the concern regarding effects of non-ASD-specific factors on ASD-symptom measures, several studies have shown strong associations between the SRS and measures of behavior problems in clinical samples of children with ASD and other psychiatric diagnoses (Bölte, Poustka, & Constantino, 2008; Charman et al., 2007; Constantino et al., 2000; Kanne, Abbacchi, & Constantino, 2009). The strength of these associations was similar for children with ASD and children with other non-ASD diagnoses (Constantino et al., 2000). In an epidemiological sample of twins, Constantino et al. (2003b) reported that scores from the Child Behavior Checklist (CBCL; Achenbach & Rescorla, 2001), a parent-report measure of psychiatric symptoms, explained 43–52% of the variance in SRS scores, although they emphasized that an additional 44% of variance was independent of behaviors captured on the CBCL. Similarly, Charman et al. (2007) reported decreased specificity of the SRS, and two other ASD screening instruments, for children with elevated behavior problems. Although many children with ASD may have additional behavior problems (Kanne et al., 2009), it is possible that associations between the SRS and measures of behavior problems reflect nonspecific difficulties rather than (or in addition to) ASD-related variation in behavior. If this were true, children with severe ASD-related impairments may not be quantitatively distinct from children with comorbid behavioral conditions, and labeling the SRS as a measure of *autism* severity could be misleading. Instead, SRS scores may be more appropriately interpreted as reflecting a broad range of impairments beyond ASD. This is of particular concern, considering that the SRS is widely used to describe the severity of ASD symptoms and/or of ASD-related social impairment in both clinical and research settings (e.g., Constantino et al., 2006; Duvall et al., 2007; Kanne et al., 2009).

Fewer studies have examined the relationship between SRS scores and factors that influence other measures of ASD-symptoms, such as age, language, and cognitive level. Although there is some evidence

that SRS scores may be influenced by these child characteristics, this is not widely acknowledged, possibly because the focus of these studies has not been to systematically examine the effects of child characteristics on SRS scores. For example, in a small clinical sample, when children were grouped by language level, nonverbal children with autism had higher scores, and their distribution was clearly differentiated from that of verbal children with autism (Constantino et al., 2000). In a larger study of families of children with ASD (Constantino, Zhang, Frazier, Abbacchi, & Law, 2010), there was a modest effect of nonverbal status (or parent-reported intellectual disability) on gender-normed SRS-T for children with ASD.

With regard to age, two studies reported that SRS scores were not significantly correlated with age in normative or clinical samples (Bölte et al., 2008; Constantino & Gruber, 2005). However, factor loadings for SRS items differed when subsets of 4 to 7-year-old and 8 to 14-year-old school children were analyzed separately (Constantino et al., 2000). Three studies including children with ASD and non-ASD diagnoses indicated nonsignificant associations with IQ (Charman et al., 2007; Constantino et al., 2003a,b, 2006), but three additional studies have reported negative correlations between SRS and FSIQ or NVIQ (Bölte et al., 2008; Constantino et al., 2000, 2007). Moreover, two of these studies (Bölte et al., 2008; Constantino et al., 2000) reported that correlations were stronger for children with ASD ($r = -.18$ to $-.42$) than non-ASD clinical controls ($r = -.04$ to $-.08$). These inconsistencies are difficult to interpret, perhaps because of small sample sizes (ranging from 37 to 127 in all, but Bölte et al., 2008) that have primarily included children with average intelligence.

In sum, studies consistently suggest a relationship between behavior problems and SRS scores; however, in spite of their implications for interpretability of scores, particularly when being used as indicators of ASD-specific severity, this is rarely acknowledged by researchers using the SRS as a quantitative measure. Moreover, the effects of age, language level, and IQ have been documented for other measures, but thorough understanding of the influence of these child characteristics on SRS scores has been obscured by small sample sizes and a lack of systematic analyses with ASD samples. Such understanding has critical implications for interpretation of SRS scores as a quantitative measure of ASD-symptoms. The present study seeks to address such limitations by investigating these relationships in a large sample of probands with ASD and their unaffected siblings. On the basis of previous studies, it is hypothesized that more behavior problems and greater expressive language impairment will be associated with higher SRS scores for probands and siblings, and that NVIQ will be negatively associated with SRS scores in probands (IQs are not available

for siblings). Consistent with other parent-rating measures (e.g., Corsello et al., 2007), it is predicted that SRS scores will be higher with increasing age.

Method

Participants

Participants were 2,368 probands and 1,913 unaffected siblings evaluated at 12 university-based centers from 2007 to 2011 as part of the Simons Simplex Collection (SSC), a genetic study of families with one child with ASD who does not have first-, second- or third-degree relatives with ASD. All probands met Collaborative Programs of Excellence in Autism (CPEA) criteria for a diagnosis of Autism, ASD, or Asperger Disorder. All siblings screened negative for ASD or indication of the broader phenotype. Detailed study procedures are included in the online appendix. Families were predominantly White (78%) and well-educated (61% maternal education of Bachelor's degree or higher). Sample demographics are provided in Table 1. Parents gave informed consent, approved by Institutional Review Boards at each university.

Measures

Autism symptoms. The SRS (Constantino & Gruber, 2005) is a parent-completed questionnaire; items describe a child's behavior in the past 6 months, yielding a raw total (SRS-Raw) and gender-normed T-score (SRS-T; intended to correct gender differences observed in normative samples). Although originally proposed as a continuously distributed, quantitative measure of ASD-severity, recent study has found bimodal distributions within affected and unaffected family members of children with ASD (Constantino et al., 2010; Virkud, Todd, Abbacchi, Zhang, & Constantino, 2009). The manual recommends use of SRS-Raw in research for comparability to early studies of the SRS, although several recent studies use SRS-T (e.g., Constantino et al., 2010).

The Autism Diagnostic Observation Schedule (ADOS; Lord, Rutter, DiLavore, & Risi, 1999) Calibrated

Severity Score (CSS) was chosen as an ASD-severity measure that is less influenced by child characteristics than raw totals (Gotham et al., 2009). This 10-point metric (higher scores reflecting greater ASD-severity) is derived from raw totals based on participants' ages and language levels. The ADI-R (Rutter et al., 2003) Current Behavior Algorithm total (ADI-Current; see Hus & Lord, 2012) was used as a parent-report measure of current ASD-symptoms.

Social development. The Vineland Adaptive Behavior Scales, Second Edition (Vineland-II; Sparrow, Cicchetti, & Balla, 2005) is a parent interview. Standard scores from the Socialization domain (VSOC) were used as a measure of social development available for probands and siblings to allow comparison between groups.

Behavior problems. Two forms of the CBCL (for children ages 18 months–5 years and 6–18 years) each yield T-scores for Internalizing (CBCL-I) and Externalizing (CBCL-E) domains and five overlapping Syndrome Scales (Anxious-Depressed, Withdrawn-Depressed, Somatic Complaints, Attention Problems, and Aggressive Behavior). CBCL-I and CBCL-E were used as estimates of behavior problems; Syndrome Scales were used for post hoc analyses.

Developmental level. Proband and sibling chronological ages in years were used as a continuous predictor for regression analyses. The Vineland-II Expressive Communication subdomain standard score (VEC) was chosen to provide a continuous indicator of expressive language abilities available for probands and siblings. For probands, ADOS Module was used as a categorical indicator of expressive language; Module-1 (single words or nonverbal) = 18.4%, Module-2 (simple phrases) = 22.8%, and Module-3 (complex sentences) = 58.8%. NVIQ was used to indicate proband cognitive level. VIQ was not included due to multicollinearity with NVIQ and because expressive language level was included separately. Only demographics, SRS, Vineland-II, and CBCL were available for siblings.

Table 1 Sample demographics

	Probands		Siblings	
	Males (<i>n</i> = 2,056)	Females (<i>n</i> = 312)	Males (<i>n</i> = 890)	Females (<i>n</i> = 1,023)
Age (years)	8.74 (3.32)	8.90 (3.60)	9.49 (3.71623)	9.46 (3.65)
SRS-Raw	97.56 (26.82)	99.32 (27.24)	20.53 (15.44)	17.22 (13.02)
SRS T-score	80.56 (12.83)	89.63 (15.05)	43.70 (7.39)	44.26 (7.19)
VSOC	71.54 (12.57)	70.15 (12.71)	<i>101.77 (11.93)</i>	<i>103.19 (11.29)</i>
CBCL-E	56.39 (10.7)	57.79 (10.2)	46.84 (9.81)	46.19 (9.42)
CBCL-I	60.35 (9.47)	59.96 (9.98)	<i>48.29 (10.19)</i>	<i>47.32 (9.96)</i>
VEC	<i>10.23 (3.06)</i>	<i>9.71 (3.03)</i>	16.02 (2.37)	16.39 (2.34)
ADI-Current	17.04 (7.30)	17.59 (7.80)		
ADOS-CSS	7.43 (1.68)	7.43 (1.73)		
NVIQ	85.81 (25.70)	78.12 (25.18)		

VSOC, Vineland-II Social Standard Score; CBCL, Child Behavior Checklist; I, Internalizing; E, Externalizing; VEC, Vineland-II Expressive Communication V-Score; ADOS-CSS, ADOS Calibrated Severity Score; NVIQ, Nonverbal IQ.

Values are expressed as Mean (SD).

Bold = $p < 0.01$.

Italics = $p < 0.05$ male versus female; Ns vary due to missing data.

Data analysis

Preliminary gender comparisons of SRS-Raw and SRS-T-Scores were conducted using SPSS 17.0 (SPSS Inc, Chicago, IL, USA) *t*-test. Pearson correlations were run between measures of ASD-symptoms and social development.

Linear regression models were analyzed separately for probands and siblings using SPSS REGRESSION. In Model-A, SRS-Raw was the dependent variable and predictors were entered in the following blocks to allow examination of the relative contribution of each set of variables: Demographics (gender = female vs. male; race = white vs. nonwhite; maternal education = graduate/bachelor degree vs. some college or less), Social Development (VSOC), Behavior Problems (CBCL-I, CBCL-E), and Developmental Level (age, VEC). All variables were centered at the mean. To examine effects of language level and age in the ASD sample, Model-B replaced the age-standardized VSOC and VEC with ADOS-CSS and ADOS-Module. To explore the relationship between parent-report measures of ASD-symptoms, ADI-Current was added, but entered last (due to its strong associations with age and language; Hus & Lord, 2012). Thus, Model B, predicting proband SRS-Raw, included: Demographics, ASD-symptoms (ADOS-CSS), Behavior Problems, Developmental Level (age, Module-1 vs. Module-3, Module-2 vs. Module-3, NVIQ), ADI-Current.

Post hoc analyses were conducted to better understand associations between SRS-Raw and CBCL. First, to examine whether behavior problems influence social development, a regression predicting VSOC (Model-C), was fit with the following variables: Demographics, Behavior Problems, Developmental Level (Age, VEC), and ASD-Symptoms (SRS-Raw). Next, to explore how the profile of differences between CBCL Syndrome Scales related to differences in SRS-Raw, differences between probands and siblings from the same family were computed for VSOC, CBCL Syndrome Scales, age, and VEC and used to predict proband-sibling differences in SRS-Raw (Model-D). Finally, to investigate whether externalizing behaviors significantly predicted SRS-Raw, a regression predicting SRS-Raw was run excluding CBCL-I from the predictors (Model-E); Model-E was otherwise identical to Model-B.

For all regression models, Cohen's f^2 was computed to assess the effect of each block of predictors, whereas controlling for all other variables; f^2 of .02, .15, and .35 reflect small, medium, and large effect sizes, respectively (Cohen, 1988). Regressions predicting SRS-Raw and SRS-T were nearly identical, therefore only analyses for SRS-Raw are reported below.

To visually demonstrate the effects of behavior problems on SRS-Raw, children were divided at

VSOC = 70 (two standard deviations below the standard mean of 100) into 'low' or 'high' groups; within each group, children were further divided into 'low' or 'high' groups at the CBCL-E clinical-concern cut-off of 64. SPSS ONEWAY and post hoc Tukey tests were used to compare children across the four VSOC/CBCL-E groups. To investigate the effects of controlling for CBCL scores, residuals from a model, including SRS-Raw as the dependent variable and CBCL-I and CBCL-E, as the predictors were compared across the four groups.

Power was adequate for all models fit (see online appendix). Given the large sample and multiple comparisons, significance level was set at $p \leq 001$.

Results

Preliminary analyses

As shown in Table 1, male siblings had higher SRS-Raw than female siblings; $t(1,747.36) = 5.03$, $p \leq 001$, but sibling SRS-T did not differ by gender. Male siblings also had somewhat lower VEC than female siblings; $t(1,905) = -3.47$, $p \leq 001$. In contrast to the sibling results, male probands and female probands did not differ on SRS-Raw, but male probands had lower SRS-T than female probands; $t(382.75) = -10.11$, $p \leq 001$. Male probands also had higher NVIQ; $t(2,366) = 4.94$, $p \leq 001$ than female probands. Table 2 shows correlations between SRS-Raw and measures of ASD symptoms and social development.

Predictors of SRS-Raw

As shown in Table 3, Model-A for probands and siblings explained 46% and 33% of variance in SRS-Raw, respectively. For both, more behavior problems and social impairment (i.e., higher CBCL-I and CBCL-E, lower VSOC) predicted higher SRS-Raw. In addition, higher SRS-Raw was associated with greater language impairment (i.e., lower VEC) for both groups, as was being male and younger for siblings only; these effects were small, but significant.

In Model B, ADOS-CSS was significant, but explained only 1% of variance in proband SRS-Raw (Block 2; Table S1). Behavior problems (Block 3) and developmental level (Block 4) had medium-to-large effects on SRS-Raw. In the final model, including all predictors, more behavior problems, higher age, and lower NVIQ were associated with higher SRS-Raw; ADI-Current explained an additional 9% of variance

Table 2 Correlations between SRS-Raw and child measures

	Age	VSOC	CBCL-I	CBCL-E	VEC	ADI-C	ADOS-CSS	NVIQ
Probands	.14	-.50	.48	.42	-.38	.52	.10	-.27
Siblings	-.08	-.27	.47	.43	-.22			

VSOC, Vineland-II Social Domain Standard Score; CBCL, Child Behavior Checklist-Internalizing T-Score; CBCL-E, Child Behavior Checklist Externalizing T-Score; VEC, Vineland-II Expressive V-Scale Score; ADI-Current, Autism Diagnostic Interview-Revised Current Behavior Algorithm Total; ADOS-CSS, ADOS Calibrated Severity Score; NVIQ, Nonverbal IQ.
Ns vary due to missing data; All correlations are significant ($p \leq 001$).

Table 3 Model-A: Predictors of SRS-Raw for probands and siblings

	Probands								Siblings							
	B	SE B	95% CI		r_{part}	R^2	ΔR^2	f^2	B	SE B	95% CI		r_{part}	R^2	ΔR^2	f^2
			Lower	Upper							Lower	Upper				
Constant	97.81	.41	97.02	98.61					18.74	.27	18.21	19.26				
Demographics						.01	.01	.01								
Gender	-.16	1.21	-2.53	2.21	.00				-2.15	.54	-3.21	-1.09	-.08	.03	.03	.03
Race	<i>-2.19</i>	<i>1.01</i>	<i>-4.16</i>	<i>-.22</i>	<i>-.03</i>				<i>2.04</i>	<i>.67</i>	<i>.72</i>	<i>3.36</i>	<i>.06</i>			
MatEduc	-.97	.85	-2.63	.69	-.02				<i>1.24</i>	<i>.56</i>	<i>.14</i>	<i>2.33</i>	<i>.04</i>			
Social development						.25	.24	.33								
VSOC	-.74	.05	-.84	-.64	-.22				-.14	.03	-.20	-.09	-.10	.09	.06	.07
Behavior problems						.45	.20	.36								
CBCL-E	.44	.05	.35	.53	.14				.30	.03	.23	.36	.16	.31	.22	.32
CBCL-I	.99	.05	.89	1.09	.29				.48	.03	.42	.55	.27			
Developmental level						.46	.01	.01								
Age	-.03	.13	-.29	.23	.00				-.53	.08	-.67	-.38	-.13	.33	.02	.04
VEC	-1.19	.20	-1.59	-.79	-.09				-.64	.13	-.90	-.39	-.09			

MatEduc, Maternal Education; VSOC, Vineland-II Social Standard Score; CBCL, Child Behavior Checklist; I, Internalizing; E, Externalizing; VEC, Vineland Expressive Communication V-Score.
 Bold = $p < .001$.
 Italics = $p < .05$.

in SRS-Raw after controlling for previous factors (Block 5).

Post hoc analyses Given that associations between SRS-Raw and CBCL-I and CBCL-E were equally large or larger than relationships with social development (VSOC) and ASD-symptoms (ADOS-CSS, ADI-Current), it was of interest to more closely examine the relationship between SRS-Raw and behavior problems. A summary of post hoc analyses is provided below (details are described in the online supplement).

First, one must consider the possibility that the association between SRS-Raw and behavior problems reflects true influences of behavior problems on social skills in probands and siblings. If true, CBCL-I and CBCL-E should be significant predictors of VSOC, a standardized measure of social development. As shown in Table S2 (Model-C), CBCL scores explained only 2–3% of variance in VSOC.

Associations between SRS-Raw and CBCL scores could also be explained by ASD-specific variation in

CBCL scales containing items that appear to describe core ASD-symptoms (e.g., CBCL-Withdrawn/Depressed). If true, proband-sibling differences in scores on these scales should be related to proband-sibling differences in SRS-Raw, whereas differences in other CBCL scales (e.g., CBCL-Attention) should not. In Model-D (Table S3), the best predictors of proband-sibling differences in SRS-Raw were differences in CBCL-Attention and CBCL-Withdrawn/Depressed. Associations between SRS-Raw and CBCL-Attention, CBCL-Withdrawn/Depressed and VSOC were of similar magnitude.

Next, the significance of externalizing behaviors as a predictor of SRS-Raw was tested in the absence of CBCL-I. As shown in Table S4 (Model-E), the relationship between CBCL-E and SRS-Raw was significant and as strong as the relationship between ADI-Current and SRS-Raw, after controlling for all other factors.

Finally, as shown in Figure 1A, comparisons of children divided into groups according to low/high

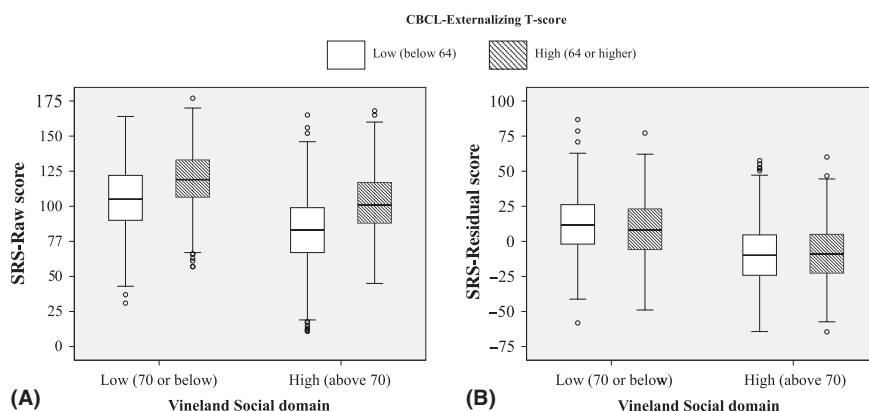


Figure 1 SRS-Raw and SRS-Residual by Vineland-Social and CBCL groups

VSOC and low/high-CBCL-E indicated significant differences in SRS-Raw, $F(3,2360) = 257.29$ $p < 001$. Tukey tests revealed that, within VSOC groups, the high-CBCL-E group had higher SRS-Raw than the low-CBCL-E group ($M_{\text{diff}} = 13.75$ and $M_{\text{diff}} = 20.31$, $p < 001$, respectively). In addition, the high-VSOC/high-CBCL-E group did not differ significantly from the low-VSOC/low-CBCL-E group ($M_{\text{diff}} = 2.16$, $p = .54$), indicating that children whose parents reported relatively good social skills and high levels of externalizing behaviors had comparable SRS-Raw to children whose parents reported relatively poor social skills and low levels of externalizing behaviors. As shown in Figure 1B, when CBCL scores were controlled, SRS-Residual scores differed significantly across the four groups; $F(3,2360) = 165.49$, $p < 001$; not surprisingly, the effects of behavior problems were diminished. Within the low-VSOC group, the low-CBCL-E group now had somewhat higher SRS-Residuals than the high-CBCL-E group ($M_{\text{diff}} = 4.10$, $p = .02$). Within the high-VSOC group, low- versus high-CBCL-E groups did not differ ($M_{\text{diff}} = -.29$, $p = .99$). In addition, the low-VSOC/low-CBCL-E group had significantly higher SRS-Residual scores than the high-VSOC/high-CBCL-E group ($M_{\text{diff}} = 20.03$, $p < 001$).

Discussion

In this study, for both probands and siblings, parent-reported behavior problems (CBCL) were strongly predictive of higher SRS-Raw, explaining similar, and often higher, proportions of variance in SRS-Raw than measures of social development (Vineland-II). For probands, SRS-Raw were also higher for older children and children with less language and lower NVIQ. When children were divided into four groups based on low or high levels of parent-reported social impairment and behavior problems, children with more externalizing behaviors had higher SRS-Raw than children with low externalizing behaviors, in spite of similar parent-reported social skills. Perhaps, most significant was the finding that children with more impaired social skills and fewer externalizing behaviors had comparable SRS-Raw to children with relatively better social skills and more externalizing behaviors. In other words, the SRS-Raw of children with good social skills, but high levels of behavior problems were indistinguishable from SRS-Raw of children with poor social skills and fewer behavior problems. It was possible to minimize these effects using SRS-Residuals from the regression model controlling for CBCL scores. These findings demonstrate that SRS-Raw are strongly influenced by non-ASD-specific child characteristics, such as internalizing and externalizing behavior problems and developmental level, highlighting the need to exercise caution when using the SRS as a continuous measure of ASD-severity or social deficits.

Associations between behavior problems, age, language, and SRS-Raw are not surprising and have been reported for several other diagnostic measures. These factors contribute to the phenotypic heterogeneity in ASD. The extent to which elevated scores on measures of behavior problems indicate distinct, co-morbid disorders or reflect secondary impairments related to ASD remains unclear (Constantino, 2011; Georgiades et al., 2010). One possibility is that these associations could be limited to parent report questionnaires. The weaker relationship between SRS-Raw and ADOS-CSS compared with that observed between SRS-Raw and ADI-Current scores highlights that method variance (i.e., clinician observation vs. parent report) may be an important factor in the measurement of ASD symptoms. However, the differential relationships between associations between SRS and VSOC for siblings and probands suggest that findings cannot be entirely attributed to parent-report bias.

Another possibility is that these associations reflect a high prevalence of behavior problems in children with ASD. Nevertheless, if the relationship between CBCL and SRS scores was explained by core ASD-features, we would not necessarily expect to find the same relationship between SRS and CBCL scores in siblings. In this study, the association between sibling SRS and CBCL was of similar magnitude to that observed for probands. In addition, although CBCL scores explained 20–26% of variance in SRS-Raw, they explained only 2–3% of variance in sibling and proband Vineland-II Social scores in this study and 2–4% of variance in proband ADOS-CSS and ADI-Current scores in a related study (Hus & Lord, 2012).

When examining the association between SRS and CBCL scores more closely, the strongest predictors of proband-sibling differences in SRS-Raw were differences in CBCL-Withdrawn/Depressed and CBCL-Attention scores. Although the Withdrawn/Depressed scale may reflect some ASD-symptoms, the CBCL-Attention scale does not include items describing core-ASD-features. Grzadzinski et al. (2011) reported that children with ADHD who score highly on SRS items related to DSM-IV criteria for ASD also score highly on items not specifically related to ASD criteria. Moreover, when internalizing symptoms were excluded from the model, externalizing behaviors had a medium-sized effect on SRS-Raw. This further suggests that the association with behavior problems cannot be entirely explained by items, which may be capturing ASD-symptoms. Although the externalizing domain is comprised of items measuring aggression, attention problems and rule-breaking behaviors, which may frequently co-occur with ASD, they are not part of the core-ASD-symptoms as defined by diagnostic criteria.

Taken together, these results indicate that SRS scores are highly influenced by behavior problems. It

remains unclear whether this is because items intended to capture social impairments (e.g., poor eye contact, difficulty with peers) lack diagnostic specificity, or whether parents interpret questions as describing qualitatively different behaviors than the ASD-symptoms that items were intended to assess (Veenstra-VanderWeele & Warren, 2012). Thus, it may be appropriate to interpret SRS scores as reflecting parents' perception of their child's overall level of impairment (which may be influenced by developmental difficulties and behavior problems, as well as ASD-symptoms), rather than as a measure of severity of core-ASD-features. This is particularly important for researchers using SRS scores as a quantitative phenotype (e.g., Duvall et al., 2007) because biological mechanisms associated with these scores may actually be markers for general impairment rather than social or ASD-specific impairments. Constantino and colleagues highlighted this idea, saying, 'Endophenotypes can be misleading ... if they do not represent truly independent subdomains of autism' (Constantino et al., 2004, p. 719). As shown in Figure 1B, one way to increase the probability that associations between SRS scores and biological mechanisms are due to ASD-related behaviors is to statistically control for non-ASD-specific influences (e.g., CBCL).

It is also important to note that although gender-normed T-scores are available to correct gender differences observed in normative samples (Constantino & Gruber, 2005), the effects of gender were minimal in our unaffected sibling sample and there were no gender differences in SRS-Raw for probands. While SRS-T 'corrects' for sibling gender differences, the same adjustment results in female probands having higher (i.e., worse) SRS-T than male probands. The only other gender difference observed for probands was that female probands were more cognitively impaired than male probands. In a recent study, gender was reportedly a significant predictor of SRS-T for both probands and unaffected siblings (Constantino et al., 2010). However, the absence of gender differences on other proband measures of ASD-symptoms suggests that using gender-normed T-scores in clinical populations may exaggerate difficulties in females. Alternatively, failing to control for gender differences using SRS-Raw could overestimate impairments in unaffected males and erroneously lead to conclusions that unaffected males are more impaired than unaffected females. Additional data are needed to better understand whether raw or T-scores are more appropriate and to inform development of standard expectations for reporting SRS scores in different sample types. Nevertheless, the effects of gender on SRS-Raw were small compared to associations with behavior problems, age, and language level. Thus, standardizing SRS-Raw to account for behavior problems and developmental level may be more crucial to interpretation of the SRS than gender-based T-scores.

Limitations

In Model A, the Vineland-II was used as a measure of social development available for both probands and siblings (because siblings were not administered the full battery of tests). The Vineland-II has demonstrated acceptable reliability and validity in normative samples, and score profiles in children with ASD reflect expected impairments in the Communication and Socialization domains (Sparrow et al., 2005; Kanne et al., 2011), suggesting that this was an appropriate measure of social impairment for both groups. However, the restricted range of VSOC scores in this sibling sample may have contributed to the relatively weak association between SRS-Raw and VSOC for siblings.

It is possible that rater contrast effects (i.e., parents comparing the proband and unaffected sibling) affected post hoc analyses of how differences in CBCL scores predicted differences in SRS-Raw. However, if this were the case, we would expect differences in SRS-Raw to predict differences on all scales or only scales assessing potentially ASD-specific symptoms (e.g., CBCL-Social Problems), which was not seen herein.

Finally, many factors known to influence parental ratings of behavior were not measured, such as parent stress levels or previous knowledge of their child's diagnosis. Although demographics, such as maternal education and race, did not emerge as significant predictors, the limited variability in our sample with respect to these characteristics limits our ability to interpret such findings. Attention to how informant characteristics influence scores on parent-report may provide important insight about ways to improve questionnaires for use as indicators of ASD severity. Notably, these issues are not limited to the SRS and should be considered for other parent-report measures being used as measures of ASD-severity. Moreover, how child and informant characteristics influence sensitivity and specificity of the SRS should be examined as they have been for other screening instruments (e.g., Corsello et al., 2007); this was not feasible in this study given the stringent criteria inclusion/exclusion criteria used for the SSC.

Conclusions

SRS scores were strongly associated with behavior problems for children with ASD and their unaffected siblings. Effects of age and expressive language level were smaller, but significant. When used as a quantitative phenotype measure in samples of children with ASD, the SRS may exaggerate impairments in children who are older or have greater behavior problems or cognitive delays. These results caution against interpretation of SRS scores as a measure of social impairment or ASD-specific severity without careful consideration of the effects of behavior problems, age, language and cognitive

level. Additional studies utilizing clinical samples of children with non-ASD diagnoses are needed to explore how these factors influence the SRS' sensitivity and specificity, as well as to inform standardization of scores for use as continuous measures of ASD-severity.

Supporting information

Additional Supporting Information may be found in the online version of this article.

Appendix S1 Study procedures.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

Acknowledgements

This research was supported by graduate fellowships from the Simons Foundation and Autism Speaks to VH,

NICHD grant R01HD065277 to SB, and Simons Foundation and NIMH grants R01MH081873 and RC1MH089721 to CL. We gratefully acknowledge Andrew Pickles for statistical consultation, Nicole Saghy for help preparing this manuscript, and the SSC families and principal investigators (A. Beaudet, R. Bernier, J. Constantino, E. Cook, E. Fombonne, D. Geschwind, D. Grice, A. Klin, D. Ledbetter, C. Martin, D. Martin, R. Maxim, J. Miles, O. Ousley, B. Peterson, J. Piggot, C. Saulnier, M. State, W. Stone, J. Sutcliffe, C. Walsh, E. Wijsman). We appreciate obtaining access to phenotypic data on SFARI Base. Approved researchers can obtain the SSC dataset described in this study ([https://ordering.base.sfari.org/~browse_collection/archive\[sfari_collection_v12\]/ui:-view\(\)](https://ordering.base.sfari.org/~browse_collection/archive[sfari_collection_v12]/ui:-view())) by applying at <https://base.sfari.org>.

Correspondence

Vanessa Hus, 530 Church Street, Ann Arbor, MI 48109-1043, USA; Email: vhus@umich.edu

Key Points

- For children with autism spectrum disorders, higher SRS scores were associated with having more behavior problems, higher age, greater impairments in expressive language, and lower nonverbal IQ.
- For unaffected siblings, having more behavior problems was strongly associated with higher SRS scores; gender, age, and expressive language also had small, but significant effects.
- To appropriately interpret SRS raw scores as an indicator of ASD-severity or ASD-related social impairment, non-ASD specific factors must be taken into account.
- Additional research is needed to consider how these factors influence diagnostic validity of the SRS in large, clinical samples not ascertained exclusively for ASD research.

References

- Achenbach, T.M., & Rescorla, L. (2001). *Manual for the ASEBA school-age forms & profiles*. Burlington, VT: ASEBA.
- Aylward, G. P. (2009). Developmental screening and assessment: What are we thinking? *Journal of Developmental & Behavioral Pediatrics, 30*, 169–173.
- Bölte, S., Poustka, F., & Constantino, J.N. (2008). Assessing autistic traits: Cross-cultural validation of the Social Responsiveness Scale (SRS). *Autism Research, 1*, 354–363.
- 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*, 554–559.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Psychology Press.
- Constantino, J.N. (2011). The quantitative nature of autistic social impairment. *Pediatric Research, 69*, 55R–62R.
- Constantino, J.N., Davis, S., Todd, R.D., Schindler, M.K., Gross, M., Brophy, S.L., ... & Reich, W. (2003a). Validation of a brief quantitative measure of autistic traits: Comparison of the Social Responsiveness Scale with the Autism Diagnostic Interview-Revised. *Journal of Autism and Developmental Disorders, 33*, 427–433.
- Constantino, J.N., & Gruber, C. (2005). *The Social Responsiveness Scale*. Los Angeles, CA: Western Psychological Services.
- Constantino, J.N., Gruber, C.P., Davis, S., Hayes, S., Passanante, N., & Przybeck, T. (2004). The factor structure of autistic traits. *Journal of Child Psychology and Psychiatry, 45*, 719–726.
- Constantino, J.N., Hudziak, J.J., & Todd, R.D. (2003b). Deficits in reciprocal social behavior in male twins: Evidence for a genetically independent domain of psychopathology. *Journal of the American Academy of Child and Adolescent Psychiatry, 42*, 458–467.
- Constantino, J.N., Lajonchere, C., Lutz, M., Gray, T., Abbacchi, A., McKenna, K., ... & Todd, R.D. (2006). Autistic social impairment in the siblings of children with pervasive developmental disorders. *American Journal of Psychiatry, 163*, 294–296.
- Constantino, J.N., Lavesser, P.D., Zhang, Y., Abbacchi, A. M., Gray, T., & Todd, R.D. (2007). Rapid quantitative assessment of autistic social impairment by classroom teachers. *Journal of the American Academy of Child and Adolescent Psychiatry, 46*, 1668–1676.
- Constantino, J.N., Przybeck, T., Friesen, D., & Todd, R.D. (2000). Reciprocal social behavior in children with and without pervasive developmental disorders. *Journal of Developmental and Behavioral Pediatrics, 21*, 2–11.
- Constantino, J.N., & Todd, R.D. (2003). Autistic traits in the general population: A twin study. *Archives of General Psychiatry, 60*, 524–530.
- Constantino, J.N., & Todd, R.D. (2005). Intergenerational transmission of subthreshold autistic traits in the general population. *Biological Psychiatry, 57*, 655–660.
- Constantino, J.N., Zhang, Y., Frazier, T., Abbacchi, A.M., & Law, P. (2010). Sibling recurrence and the genetic epidemiology of autism. *American Journal of Psychiatry, 167*, 1349–1356.
- Corseello, C., Hus, V., Pickles, A., Risi, S., Cook, E.H., Leventhal, B.L., & Lord, C. (2007). Between a ROC and a hard place:

- Decision making and making decisions about using the SCQ. *Journal of Child Psychology and Psychiatry*, 48, 932–940.
- Duvall, J.A., Lu, A., Cantor, R.M., Todd, R.D., Constantino, J.N., & Geschwind, D.H. (2007). A quantitative trait locus analysis of social responsiveness in multiplex autism families. *American Journal of Psychiatry*, 164, 656–662.
- Georgiades, S., Szatmari, P., Duku, E., Zwaigenbaum, L., Bryson, S., Roberts, W., ... & Pathways in ASD Study Team. (2010). Phenotypic overlap between core diagnostic features and emotional/behavioral problems in preschool children with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, 41, 1321–1329.
- Gotham, K., Pickles, A., & Lord, C. (2009). Standardizing ADOS scores for a measure of severity in autism spectrum disorders. *Journal of Autism and Developmental Disorders*, 39, 693–705.
- Grzadzinski, R., Di Martino, A., Brady, E., Mairena, M. A., O’Neale, M., Petkova, E., ... & Castellanos, F. X. (2011). Examining autistic traits in children with ADHD: Does the autism spectrum extend to ADHD? *Journal of Autism and Developmental Disorders*, 41, 1178–1191.
- Hus, V., & Lord, C. (2012). Use of the Autism Diagnostic Interview-Revised as a measure of ASD severity. *Journal of Autism and Developmental Disorders*. Advanced online publication. doi: 10.1007/s10803-012-1576-y.
- Kanne, S.M., Abbacchi, A.M., & Constantino, J.N. (2009). Multi-informant ratings of psychiatric symptom severity in children with autism spectrum disorders: The importance of environmental context. *Journal of Autism and Developmental Disorders*, 39, 856–864.
- Kanne, S.M., Gerber, A.J., Quirnbach, L.M., Sparrow, S.S., Cicchetti, D.V., & Saulnier, C.A. (2011). The role of adaptive behavior in autism spectrum disorders: Implications for functional outcome. *Journal of Autism and Developmental Disorders*, 41, 1007–1018.
- Lord, C., Rutter, M., DiLavore, P.S., & Risi, S. (1999). *Autism diagnostic observation schedule (ADOS)*. Los Angeles, CA: Western Psychological Services.
- Mayes, S.D., & Calhoun, S.L. (2010). Impact of IQ, age, SES, gender, and race on autistic symptoms. *Research in Autism Spectrum Disorders*, 5, 749–757.
- Rea, L.M., & Parker, R.A. (1992). *Designing and conducting survey research: A comprehensive guide*. San Francisco: Jossey-Bass.
- Rutter, M., Le Couteur, A., & Lord, C. (2003). *Autism diagnostic interview-revised*. Los Angeles, CA: Western Psychological Services.
- Sparrow, S.S., Cicchetti, D.V., & Balla, D.A. (2005). *Vineland Adaptive Behavior Scales* (2nd edn). America: Circle Pines, MN.
- Veenstra-VanderWeele, J., & Warren, Z. (2011). Social communication deficits in the general population: How far out does the autism spectrum go? *Journal of the American Academy of Child and Adolescent Psychiatry*, 50, 326–328.
- Virkud, Y.V., Todd, R.D., Abbacchi, A.M., Zhang, Y., & Constantino, J.N. (2009). Familial aggregation of quantitative autistic traits in multiplex versus simplex autism. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, 150, 328–334.
- Warren, Z., Vehorn, A., Dohrmann, E., Nicholson, A., Sutcliffe, J.S., & Veenstra-VanderWeele, J. (2012). Accuracy of phenotyping children with autism based on parent report: What specifically do we gain phenotyping “rapidly”? *Autism Research*, 5, 31–38.

Accepted for publication: 13 June 2012

Published online: 24 July 2012