Research in Autism Spectrum Disorders 95 (2022) 101973

Contents lists available at ScienceDirect

Research in Autism Spectrum Disorders

journal homepage: www.elsevier.com/locate/rasd

The psychometric properties of the Norwegian version of the social responsiveness scale in a neuropediatric sample

Sabine Kaiser [a,2], Kjersti Åby Bergquist [b], Marianne Berg Halvorsen [c,*,1]

[a] Regional Center for Child and Youth Mental Health - North, UiT The Arctic University of Norway, Tromsø, Norway
[b] Department of Psychology, UiT The Arctic University of Norway, Tromsø, Norway
[c] Department of Pediatric Rehabilitation, University Hospital of North Norway, Tromsø, Norway

ARTICLE INFO

Keywords:
Assessment
Autism spectrum disorder
Neurodevelopmental disorder
Psychometrics
Social responsiveness scale

ABSTRACT

Background: The current study is an examination of the psychometric properties of the Norwegian Social Responsiveness Scale (SRS), a measure of deficits in social behavior, in a neuropediatric outpatient sample of children and adolescents with neurological and neurodevelopmental disorders.
Method: The internal consistency of the SRS, the convergent validity of the SRS with the Vineland Adaptive Behavior Scale-II (VABS-II), the Strengths and Difficulties Questionnaire (SDQ), and the Aberrant Behavior Checklist (ABC) were examined, in addition to four different factor models of the SRS (i.e., a one-factor, the original five-factor, a second-order five-factor model, and a 16-item one-factor model) using confirmatory factor analyses.
Results: There was satisfactory internal consistency on all subscales, except for the Social Awareness subscale. The SRS showed a somewhat meaningful overlap with parts of the related scales on the VABS-II, the SDQ, and the ABC. Model fit indices were mixed for evaluating the four different factor models. Overall, however, the model fit was rather poor.
Conclusions: The original SRS subscales showed adequate internal consistency and satisfactory convergent validity on some of the subscales. The construct validity in terms of factor structure was not acceptable. Future research should examine the psychometric properties of an improved version of the SRS, especially in terms of improving the scale's construct validity.

1. Introduction

Autism spectrum disorder (ASD) is a neurodevelopmental condition characterized by qualitative abnormalities in relation to reciprocal social interaction, such as restricted communication in addition to repetitive patterns of activities and interests (World Health Organization, 2018). There is great individual variation in how the core symptoms develop and in their severity. In Norway, 0.7% of children receive an ASD diagnosis before the age of eight (Surén et al., 2019). The diagnostic understanding of autism has changed over the years; currently, the autism spectrum is included in the newest version of the Diagnostic and Statistical Manual of Mental Disorders (DSM-5; American Psychiatric Association, 2013) and the International Classification of Diseases (ICD-11; World

* Correspondence to: Department of Pediatric Rehabilitation, University Hospital of North Norway, P.O. Box 100, N-9038 Tromsø, Norway.
E-mail address: marianne.berg.halvorsen@unn.no (M.B. Halvorsen).
[1] ORCID: 0000–0002-5962–7612
[2] ORCID: 0000-0002-2081-7734

https://doi.org/10.1016/j.rasd.2022.101973
Received 13 December 2021; Received in revised form 6 April 2022; Accepted 9 April 2022
Available online 29 April 2022

Health Organization, 2018). This change allows for an assessment of an individual's challenges on the autism spectrum in terms of severity of symptoms rather than fitting into categories that may not completely encompass the true nature of the phenomenon.

ASD diagnosis is established on the basis of behavior presentation, typically using the "gold standard" assessment tools: the Autism Diagnostic Observation Schedule (ADOS; Lord et al., 2000; Lord et al., 2012) and the Autism Diagnostic Interview (ADI-R; Rutter et al., 2003). However, there are several screening tools used to aid in the diagnostic process, including the Social Responsiveness Scale (SRS; Constantino & Gruber, 2012), which is frequently used. In line with the current understanding of ASD diagnostics, the SRS was developed to provide a clinical severity measure of a varied range of deficits in reciprocal social behaviors; it offers an overall scale and comprises the five subscales of social awareness, social cognition, social communication, social motivation, and autistic mannerisms.

The internal consistency of the subscales in terms of Cronbach's alpha varies among study groups (Wang et al., 2012) and among studies and has been found unacceptable (<0.70) for three subscales (social awareness, social cognition, social motivation) in Cen et al. (2017), for one subscale (social awareness) in Duku et al. (2013) and Wigham et al. (2012), respectively, and for two subscales (social cognition, social motivation) in Shahrivar et al. (2020). Fombonne et al. (2012), on the other hand, found acceptable to excellent Cronbach's alpha for the five scales in his study.

In addition, Bolte et al. (2008) reported negative correlations with the Vineland Adaptive Behavior Scale (VABS-II; e.g., SRS and VABS-II total score $r = -0.36$ and SRS and VABS-II Socialization scale $r = -0.41$) and concluded that construct validity was ensured by these correlations. Furthermore, Wigham et al. (2012) stated that convergent validity was supported through correlations between the SRS total score and the Strengths and Difficulties Questionnaire (SDQ; Goodman, 1997) total problem score ($r = 0.70$, $p < .01$) in addition to a significant, negative correlation between the SRS total score and the SDQ pro-social scale ($r = -0.45$, $p < 01$).

Given that the SRS appears in several languages, it is also important to establish linguistic equivalence and to determine the consistency of the factor structure whenever possible. The factor structure of the SRS has been examined in several exploratory factor analyses (EFA; Duku et al., 2013; Gau et al., 2013; Moul et al., 2015; Sturm et al., 2017) and confirmatory factor analyses (CFA; Cen et al., 2017; Duku et al., 2013). The original five-factor model seems to have poor goodness-of-fit (Cen et al., 2017; Duku et al., 2013), as well as a second-order five-factor model and a one-factor model (Duku et al., 2013). Based on the EFA and CFA results, suggestions have been made to change the SRS by removing items or changing the number of factors (Duku et al., 2013; Gau et al., 2013; Moul et al., 2015; Sturm et al., 2017). Sturm et al., (2017) developed and validated a 16-item SRS form and concluded that this short form "may possess superior psychometric properties compared to the original scale" (p. 1053) in terms of reliability, relatedness to the original SRS full version, and relation with gold standard assessment tools.

However, Constantino and Gruber (2005) originally described the five subscales of the SRS as 'treatment subscales', as they are based on clinical experience and not derived empirically by factor analysis. The decision to continue to use the subscales was based on the assumption that the subscales are clinically relevant and measure different aspects of social reciprocity, particularly in regard to ASD, rather than a unidimensional approach (Constantino & Gruber, 2005). Although the subscales were not intended to be used as independent measures, in research praxis they often are. Additionally, only one previous CFA using the non-English version was conducted in a Chinese sample (Cen et al., 2017). None of these studies have been conducted using the Norwegian or another Scandinavian version of the SRS, indicating the need for further research.

1.1. Study aim

The aims of the study were to examine the psychometric properties of the SRS in a relatively large neuropediatric sample of children with neurodevelopmental and neurological disorders. First, we examine the characteristics of the different diagnostic groups and the total sample in addition to the internal consistency (i.e., Cronbach's alpha) for the five SRS subscales. Second, we examine the convergent validity by correlating the SRS and its five subscales with other relevant measures addressing similar concepts (i.e., the VABS-II (Sparrow et al., 2011), the Aberrant Behavior Checklist [ABC] (Aman & Singh, 1994), and the SDQ (Goodman, 1997)). It was hypothesized that there would be (i) negative correlations between the SRS subscales and the VABS-II domains Communication and Socialization and the VABS-II total scale, (ii) positive correlations between SRS subscales and total scale and the ABC subscales Social Withdrawal, Inappropriate Speech, and Stereotyped Behavior; (iii) positive correlations between the SRS subscales and total scale and the SDQ subscale Peer Problems; and (iiii) a negative correlation between the SRS total scale and the SDQ Prosocial subscale. We did not specify the expected strength of the correlations a priori. This was because the measures/subscales we compared to those of the SRS were used complementarily and not interchangeably with the SRS. Third, we conducted four CFAs of the SRS. Because the five subscales are considered 'treatment subscales' (Constantino & Gruber, 2005) and used in research praxis, we examined a one-factor model, a five-factor model, and a second-order five-factor model (Cen et al., 2017; Duku et al., 2013). However, these factor models seem to have poor fit (Cen et al., 2017; Duku et al., 2013), highlighting a need to further elucidate the underlying factor structure of the SRS. Therefore, we also examined the short version of the SRS consisting of 16 items that load on one factor (Sturm et al., 2017). Because of previously reported findings, we expected the short version of the SRS to have the best model fit compared to the other remaining models (Cen et al., 2017; pp. 1, 1273; Duku et al., 2013; Sturm et al., 2017).

2. Methods

2.1. Participants

The sample included a total of 343 patients (114 girls and 229 boys) referred for developmental/neurological assessment to two neuropediatric outpatient clinics in northern Norway. The majority were referred to neuropediatric outpatient clinics in specialist

health services at the University Hospital of North Norway (UNN; $n = 292$, 85.1%), and the remaining patients were referred to the Finnmark Hospital Trust ($n = 51$, 14.9%). They were referred by general practitioners ($n = 234$, 68.2%) or medical specialists in specialist health services ($n = 109$, 31.8%). The neuropediatric outpatient clinics provide interdisciplinary assessment with special education therapists, pediatricians, physiotherapists, and neuropsychologists. The exclusion criteria for this study included an age below four years. This was due to a lack of suitability of one or more instruments for that age group. In addition, a lack of parental fluency in Norwegian was also a criterion for exclusion. Based on the criteria in the SRS manual, participants with incomplete SRS forms (i.e., forms with more than the minimum of 16 items missing) were also excluded ($n = 22$).

The children were aged 4–18 years ($M = 10.11$, $SD = 3.85$). The mean index score of the individualized standardized intelligence tests was 79.37 ($SD = 17.23$, range = 40–140). Over half of the children (68%) had one parent with a college or university degree. The most frequent neurodevelopmental disorders in the sample, in descending order, were (a) specific developmental disorders (32%), (b) ID (21%), (c) other diseases of the nervous system, such as epilepsy and cerebral palsy (16%), (d) ASD (15%), (e) ADHD (14%), and (f) congenital malformations and chromosomal abnormalities (13%). The diagnoses were not mutually exclusive, so a given child could have more than one diagnosis. For a further description of the design and sample, see Halvorsen et al. (2019).

2.2. Measures

The Norwegian SRS (Constantino & Gruber, 2005) is a 65-item questionnaire for children between 4 and 18 years of age used to assess the severity of social impairment related to ASD and measure changes in the severity of social impairment as a response to an intervention or over time. The SRS version used in the present study is identical to the newer SRS-II school age form (Constantino & Gruber, 2012). It consists of five subscales: Social awareness (e.g., "Walks in between two people who are talking", 8 items), Social cognition (e.g., "Takes things too literally and doesn't get the real meaning of the conversation", 12 items), Social communication (e.g., "Talks with an unusual tone of voice (e.g., talks like a robot or like he or she is giving a lecture)", 22 items), Social motivation (e.g., "Seems much more fidgety in social situations than when alone", 11 items), and Autistic mannerisms (e.g., "Has repetitive, odd behaviors such as hand flapping or rocking", 12 items). The SRS is rated by a parent/caretaker or a teacher. The SRS uses a Likert scale with four response categories: 0 = "not true", 1 = "sometimes true", 2 = "often true" and 3 = "almost always true". The total raw scores can range from 0 to 195, with a higher score indicating a higher severity of symptoms. Psychometric properties of the SRS have been reported in the introduction.

The VABS-II (Sparrow et al., 2011) is a semistructured interview and a widely used assessment tool for the evaluation of adaptive behavior in 2–21 year olds (Sparrow et al., 2011). It evaluates adaptive behavior by inquiring about an individual's ability to manage oneself, personally and socially, in daily activities. Adaptive behavior is evaluated in four domains: communication, daily skills, social abilities, and motor skills (only for children younger than seven years). The psychometric properties of the Norwegian version of the VABS-II have been described in the manual using a representative Scandinavian sample of 1673 parents of children ages 2–21, indicating good psychometric properties, and Scandinavian norms are available (Sparrow et al., 2011). The construct validity of the instruments is among others supported through CFAs. Split-half reliability was satisfactory for the VABS-II Total (Sparrow et al., 2011).

The Norwegian ABC (Aman & Singh, 1994) is a behavior rating scale used to assess behavior problems in children and adults with developmental disabilities. It is typically filled out by a parent or other close person in approximately 15 minutes and uses a 4-point Likert scale, from 0 ("no problem") to 3 ("serious problem"). It includes 58 items and consists of five subscales: Irritability (15 items), Passivity and Social Withdrawal (16 items), Stereotypical Behavior (7 items), Hyperactivity (16 items), and Inappropriate Speech (4 items). The Norwegian version of the ABC has been found to have satisfactory psychometric properties in the current sample in terms of internal consistency and correlations with other relevant constructs (Halvorsen et al., 2019).

The Norwegian SDQ parent report (Goodman, 1997) is a mental health questionnaire designed for children aged 4–16. It consists of 25 items divided into five subscales with five items each. The first four are categorized as problem scales: emotional symptoms, hyperactivity-inattention, conduct problems, and peer problems. The final subscale assesses prosocial behavior. The SDQ uses a Likert scale, from 0 ("not true") to 2 ("certainly true"). The psychometric properties of the Norwegian SDQ have been examined in the current sample (Kaiser & Halvorsen, 2022). For the SDQ parent version, internal consistency was adequate and to good for the four scales used in the current study (i.e., emotional symptoms, hyperactivity-inattention, peer problems, and prosocial behavior scales). Correlations with other constructs, such as the ABC scales, were meaningful. The results of the CFAs indicated the best model fit for the parent version compared to the self-report or teacher version (Kaiser & Halvorsen, 2022).

Intellectual Function: Intellectual functioning/level was defined by the Full Scale IQ (FSIQ), which was assessed through an age-appropriate standardized Wechsler intelligence test (Wechsler, 2007, 2008a, 2008b, 2009, 2012). Due to insufficient completion of subtests on the Wechsler test, the FSIQ score for a small number of children was assessed using Raven's Colored Progressive Matrices (n = 14; Raven, 2004). The psychometric properties of the Norwegian versions of Wechsler Intelligence tests have been described in the manuals. The Norwegian version of the WISC-IV was adjusted to the Norwegian context and has Norwegian norms. The norms were based on a Norwegian standardization selection of 418 children aged 6 years to 16 years and 11 months. The Norwegian version has proven to have good reliability, with coefficients between 0.86 and 0.98 (Wechsler, 2009). The validity of the Norwegian version examined by confirmatory factor analysis showed that the original WISC-IV factor structure was transferable to the Norwegian version (Wechsler, 2009).

2.3. Procedure

The interdisciplinary assessments at the neuropediatric clinics typically lasted two days. Assessments of the different aspects of

neurodevelopmental disorders were completed by neuropsychologists, clinical psychologists, pediatricians, and physiotherapists. All children completed standardized intelligence tests and the VABS-II with a neuropsychologist or clinical psychologist. Assessments of neurological/neurodevelopmental disorders were completed by a pediatrician/neurology specialist with the use of instruments such as EEG, MRI caput and genetic testing. Assessments regarding motor delays and muscle disease were completed by a physiotherapist. Diagnoses of neurodevelopmental and neurological disorders, including ASD and ADHD, were obtained from the interdisciplinary assessment. The ICD-10 criteria were used to code diagnoses (World Health Organization, 1992). The presence of an ID was operationalized as scoring below 70 on both the standardized intelligence test and the VABS-II. The study was approved by the appropriate ethics committee. The data protection officer at UNN and Finnmark Hospital Trust approved the use of deidentified data for research purposes.

2.4. Data analysis

SPSS 26 was used to calculate the descriptive statistics and internal consistency (i.e., Cronbach's alpha). Cronbach's alpha (α) equal to or above .90 is excellent, .80 to .90 is good, .70 to .80 is adequate and below .70 is inadequate (European Federation of Psychologists' Associations, 2013). Pearson's correlations were used to address the relationship between SRS subscales, convergent validity, age, and gender. The European Federation of Psychologists' Associations (2013) guideline suggests convergent validity scores (*r*) between. 55 to .65 are adequate, .65–0.75 are good, and above .75 is excellent.

Furthermore, four confirmatory factor analyses were conducted to examine the factor structure of the Norwegian SRS using SPSS AMOS and the maximum likelihood estimator. In accordance with Constantino and Gruber (2005) and Duku et al. (2013), a one-factor model, the original five-factor model, and a second-order five-factor model were analyzed. In addition, a short version of the SRS consisting of 16 items that load on one factor was analyzed in accordance with Sturm et al. (2017). Model fit was evaluated using goodness-of-fit indices, including chi-square, chi-square/df ratio, root mean squared error of approximation (RMSEA), Tucker–Lewis index (TLI), and comparative fit index (CFI). The chi-square/df ratio was used to assess model fit with a range for acceptance from high 5.0 to low 2.0 (Hooper et al., 2008). The remaining criteria for good fit were based on Hu and Bentler (1999), with an RMSEA cutoff often set at .06, with an upper limit of .07, and for the CFI and TLI, a value equal to or greater than 0.95. The Akaike information criterion (AIC) was used to compare models. The model with the lowest value is considered the best fitting model (Hooper et al., 2008). All analyses were conducted using the raw scores of the SRS.

3. Results

3.1. Participant characteristics and SRS scores

Table 1 provides demographic and clinical data in addition to means and standard deviations for the SRS subscales for the main diagnostic groups and the total sample. The majority of the 343 patients were boys ($n = 229$, 66.8%), and the mean age was 10.11 ($SD = 3.85$). Approximately 38 children were diagnosed with ASD, 36 with ADHD, 53 with intellectual disability (ID), and the remaining children ($n = 216$) had different diagnoses, such as specific developmental disorders, epilepsy, and cerebral palsy. The mean score on

Table 1
Descriptive statistics for the different diagnostic groups and the total sample (N = 343).

	ASD	ADHD	ID	Remaining	Total
	M(SD)	*M(SD)*	*M(SD)*	*M (SD)*	*M(SD)*
N	38	36	53	216	343
Males *n* (%)	28 (73.7%)	29 (80.6%)	32 (60.4%)	140 (64.8%)	229 (66.8%)
Females *n* (%)	10 (26.3%)	7 (19.4%)	21 (39.6%)	76 (35.2%)	114 (33.2%)
Age	9.61 (3.98)	10.75 (3.43)	9.58 (4.00)	10.22 (3.86)	10.11 (3.85)
FSIQ	94.52 (12.59)	80.10 (11.78)	58.09 (7.65)	78.08 (16.22)	76.37 (17.23)
VABS-II Total score	60.24 (14.28)	62.34 (11.63)	56.43 (8.09)	72.01 (15.25)	67.13 (15.24)
Parental education					
Primary school	–	3 (8.3)	2 (3.8)	6 (2.8)	11 (3.2)
High School *n* (%)	9 (23.7%)	13 (36.1%)	14 (26.4%)	63 (29.2%)	99 (28.9%)
College/University *n* (%)	29 (76.3%)	20 (55.6%)	37 (69.8%)	147 (68.1%)	233 (67.9%)
SRS Awareness	10.35 (4)	9.94 (3.46)	9.66 (3.04)	7.77 (4.10)	8.58 (4.01)
SRS Cognition	13.74 (5.83)	14.42 (6.97)	13.10 (5.76)	10.59 (6.68)	11.72 (6.64)
SRS Communication	24.77 (10.43)	22.88 (9.84)	19.60 (8.00)	17.77 (10.62)	19.38 (10.44)
SRS Motivation	14.97 (5.73)	13.12 (6.07)	12.14 (5.45)	11.13 (5.91)	11.93 (5.94)
SRS Mannerisms	14.49 (7.27)	11.88 (7.47)	9.23 (5.67)	8.81 (6.96)	9.79 (7.08)
SRS total	76.06 (27.78)	71.93 (28.67)	63.14 (23.28)	55.34 (30.56)	60.65 (29.92)

Note. ASD = autism spectrum disorder (without ADHD or ID); ADHD = attention-deficit/hyperactivity disorder (without ASD or ID); ID = intellectual disability (without ASD or ADHD).
FSIQ = Full Scale Intelligence Quotient; VABS-II =Vineland Adaptive Behavior Scale-II score; SRS raw score = Social Responsiveness Scale. The subgroups did not reach an *N* of 343 because we only included subgroups relevant to the focus of this study. The remaining group included all participants without isolated cases of ASD, ADHD or ID ($n = 216$; i.e., specific developmental disorders, other diseases of the nervous system such as epilepsy and cerebral palsy, and congenital malformations and chromosomal abnormalities).

the SRS total was 60.65 ($SD = 29.92$) for the total sample. Cronbach's alpha was higher than .70 for four of the SRS subscales ($\alpha = 0.79$ –0.89), with the exception of the Social Awareness subscale ($\alpha = 0.68$).

3.2. The Relationship between SRS and Demographic and Clinical Variables

Table 2 includes the correlations among the examined variables as well as descriptive statistics. All subscales and the total scale of the SRS correlated significantly with VABS-II subscales and the VABS-II total scale, the ABC subscales, and the SDQ subscales. The SRS scales and the total scale correlated negatively with the VABS-II Communication subscale (ranging from $r = -0.29$ to $r = -0.50$) and with the VABS-II Socialization subscale (ranging from $r = -0.39$ to -0.63). The SRS scales and total scale correlated positively with all ABC scales (ranging from $r = 0.27$ to.72). The SRS scales and the total scale correlated positively with the SDQ Peer Problems scale (ranging from $r = 0.43$ to.57) and negatively with the SDQ Prosocial scale (ranging from $r = -0.34$ to -0.55).

3.3. Confirmatory factor analysis

Table 3 presents the results of the CFAs. The χ^2 is significant ($p < .001$) in all four models, and the χ^2 /df ratio is within the acceptable range for all models. The TLI and CFI are under the recommended cutoff of .95 in all models. The AIC indicates the best model fit for the SRS short version. The standardized factor loadings for the SRS short version are presented in Fig. 1. All factor loadings were larger than .40, except for items 7, 38, and 54 (>0.30).

4. Discussion and implications

The SRS is a widely used rating scale among people with ASD and other neurodevelopmental disorders with co-occurring autistic traits. Based on previous research, further examination of the psychometric properties of the SRS was necessary, and as far as the authors know, none of the previously conducted psychometric studies have included a Norwegian or Scandinavian population.

The internal consistency of four of the five scales of the SRS was adequate and to good. However, it should be noted that the Social Awareness subscale was the only subscale that did not have adequate internal consistency. This is in accordance with findings from several previously published studies (Cen et al., 2017; Duku et al., 2013; Wigham et al., 2012). A low Cronbach's alpha could, for example, indicate poor interrelatedness between items or that the scale measures heterogeneous constructs (Tavakol & Dennick, 2011).

Regarding the convergent validity, correlations between the SRS and the other relevant measures, that is, the VABS-II, the ABC, and the SDQ, were all significant and in the expected direction. In accordance with the guidelines from the European Federation of Psychologists' Associations (2013), good scores were found between the SRS total scale and three subscales (i.e., the SRS Motivation subscale, the SRS Communication, and the SRS Mannerism subscales) with the ABC Social withdrawal scale, indicating that similar

Table 2
Bivariate Relationship between SRS Subscales and Total Score and Demographic/Clinical Variables.

Variable	*M (SD)/n* (%)	SRS Awa	SRS Cog	SRS Com	SRS Mot	SRS Man	SRS Total
Age	10.11 (3.85)	-0.06	0.06	0.11 *	0.22 * **	0.05	0.11
Gender[a]	229 (66.8%)	-0.00	-0.03	0.04	-0.05	0.08	-0.00
FSIQ	76.37 (17.32)	-0.02	-0.23 * **	-0.03	-0.06	-0.05	-0.12
VABS-II							
Communication	64.66 (13.85)	-0.46 * **	-0.50 * **	-0.47 * **	-0.29 * **	-0.38 * **	-0.49 * **
Daily Living Skills	74.45 (14.62)	-0.46 * **	-0.46 * **	-0.37 * **	-0.22 * **	-0.41 * **	-0.44 * **
Socialization	72.33 (15.68)	-0.62 * **	-0.56 * **	-0.63 * **	-0.39 * **	-0.55 * **	-0.64 * **
Total	67.13 (15.24)	-0.57 * **	-0.58 * **	-0.55 * **	-0.36 * **	-0.52 * **	-0.60 * **
ABC							
Irritability	6.13 (7.38)	0.50 * **	0.57 * **	0.53 * **	0.42 * **	0.54 * **	0.58 * **
Social Withdrawal	5.74 (6.25)	0.46 * **	0.55 * **	0.69 * **	0.72 * **	0.62 * **	0.72 * **
Stereotype	1.52 (2.71)	0.34 * **	0.40 * **	0.44 * **	0.27 * **	0.56 * **	0.49 * **
Hyperactivity	9.76 (9.99)	0.55 * **	0.55 * **	0.53 * **	0.29 * **	0.56 * **	0.57 * **
Inappropriate Speech	1.73 (2.3)	0.45 * **	0.56 * **	0.56 * **	0.32 * **	0.62 * **	0.60 * *
SDQ							
Emotion	3.37 (2.59)	0.17 * *	0.38 * **	0.32 * **	0.47 * **	0.31 * **	0.38 * **
Inattention/hyperactivity	3.85 (2.59)	0.49 * **	0.47 * **	0.45 * **	0.22 * **	0.42 * **	0.47 * **
Peer Problems	3.85 (2.5)	0.43 * **	0.43 * **	0.57 * **	0.47 * **	0.47 * **	0.55 * **
Prosocial	7.03 (2.38)	-0.54 * **	-0.42 * **	-0.55 * **	-0.34 * **	-0.43 * **	-0.53 * **
ASD *n* (%)	38 (11.1%)	0.16 * *	0.11	0.18 * **	0.18 * **	0.23 * **	0.19 * *
ADHD *n* (%)	36 (10.5%)	0.12 *	0.14 *	0.12 *	0.07	0.10	0.13 *
ID *n* (%)	53 (15.5%)	0.11 *	0.09	0.01	0.02	-0.03	0.04

Note. [a]Gender = male. SRS = Social Responsiveness Scale; Awa = Social Awareness; Cog = Social Cognition; Com = Social Communication; Mot = Social Motivation; Man = Autistic Mannerisms. FSIQ = full scale intelligence quotient; VABS-II = Vineland Adaptive Behavior Scale-II; ABC = Aberrant Behavior Checklist; SDQ = Strengths and Difficulties Questionnaire; ASD = Autism Spectrum Disorder; ADHD = Attention-Deficit/Hyperactivity Disorder; ID = Intellectual Disability.
*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

Table 3
Results of the confirmatory factor analyses (N = 343).

	df	χ^2	χ^2/df	RMSEA (90% CI)	TLI	CFI	AIC
1. One factor	2015	5570.87[a]	2.77	.072 (0.070;0.074)	0.589	0.614	5960.87
2. Five factor	2005	5248.78[a]	2.62	.069 (0.067;0.071)	0.623	0.648	5658,78
3. Second-order five-factor	2010	5305.09[a]	2.64	.069 (0.067;0.072)	0.618	0.642	5705.09
4. SRS short version	104	394.02[a]	3.79	.088 (0.078;0.097)	0.799	0.846	490.02

Note. χ^2 = Chi-Square; df = Degrees of Freedom; χ^2/df = Normed Chi-Square; RMSEA = Root Mean Squared Error of Approximation; 90% CI = 90% Confidence Interval; TLI = Tucker–Lewis Index; CFI = Comparative Fit Index; AIC = Akaike Information Criterion.
[a] $p < 0.001$.

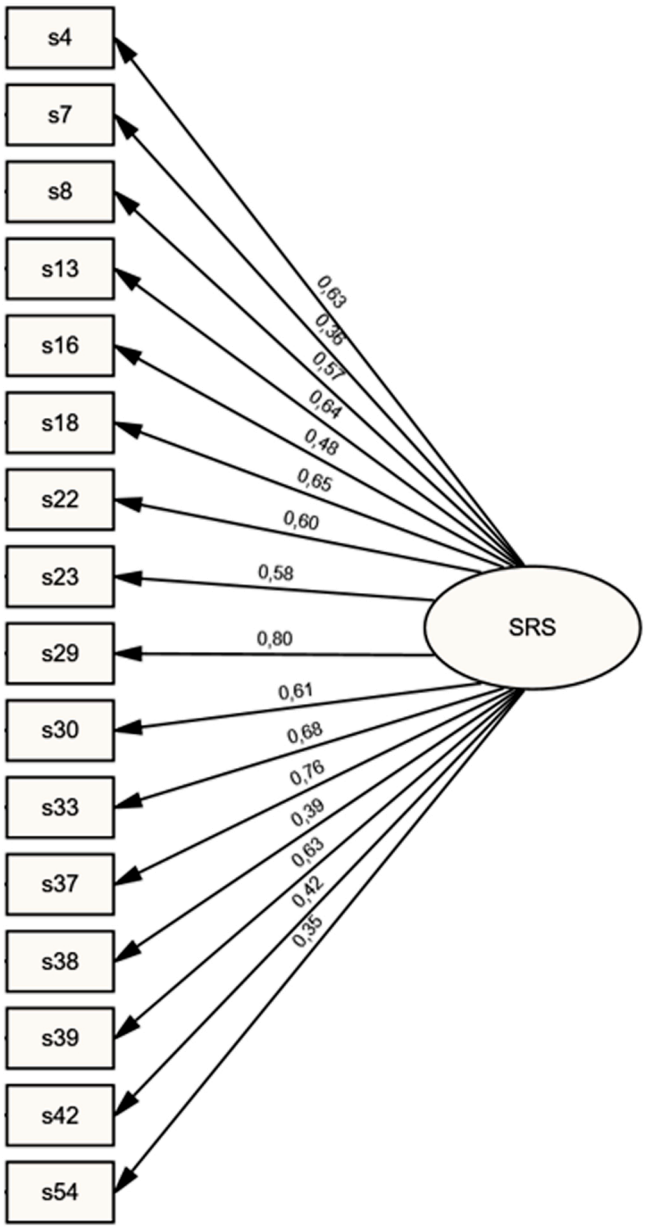

Fig. 1. Figure 1 displays the short version of the Social Responsiveness Scale with its 16 items and factor loadings (Sturm et al., 2017).

constructs have been examined. Similarly, SRS Mannerism subscales and the SRS total scale had good correlations with the ABC Inappropriate speech scale. The remaining correlations between the SRS and the ABC were lower, indicating that the constructs are less similar (European Federation of Psychologists' Associations, 2013).

Correlations between the SRS total scale and the VABS-II scales were adequate for the Socialization scale and the VABS-II total score in the current sample and in accordance with EFPA (2013) guidelines. However, findings in the literature are mixed. While Pine et al. (2006) found excellent correlations between the SRS and the VABS-II total score, Bolte et al. (2008) and Charman et al. (2007) found correlations that indicate that these instruments are less similar. Similar results to the one from Charman et al. (2007) are reported by Hus et al. (2013) for the SRS total score and the VABS-II Socialization scale.

In contrast to Wigham et al. (2012), the current study found relatively low correlations between the SRS and the SDQ; however, correlations were adequate, for example, for the SRS total score and the SDQ peer problems scale. The fact that VABS-II, ABC, and SDQ do not overlap more with SRS may be because their targets are not identical. These measures are used complementarily and not interchangeably with the SRS. However, they do have certain similarities and do correlate significantly and in the expected direction.

To evaluate the factor structure of the SRS, several CFAs were conducted (DiStefano & Hess, 2005). Model fit is generally evaluated by using a combination of different fit indices (Hu & Bentler, 1999). In the current study, the chi-square test was significant in all models, indicating that the theoretical models did not fit the data (Matsunaga, 2010). The chi-square/degrees of freedom ratios were within the limit of the cutoff value, indicating acceptable model fit. An RMSEA smaller than .06, or at least smaller than .07, and a Tli and CFI larger than .95 indicate good model-to-data fit (Hu & Bentler, 1999). In the current study, the RMSEA was on the upper limit to .07 in all models, and the TLI and CFI were by far under the recommended cutoff value of .95. Taking all model fit indices into account, this leads to the overall conclusion that the models had no acceptable model fit. In line with the current findings, previous CFAs of the full version of the SRS have shown that the five-factor model seems to have poor goodness-of-fit (Cen et al., 2017; Duku et al., 2013), as well as a second-order five-factor model and a one-factor model (Duku et al., 2013). Because of the poor model fit, it is not surprising that several of the items in the present study were found to have poor factor loadings in all models. As an example, item 43 ("Separates easily from caregiver"; reverse scored) was found to be nonsignificant in the current as well as in other studies (Bolte et al., 2008; Cheon et al., 2016). One possibility to improve model fit is to remove items with small factor loadings from the analysis (Hooper et al., 2008). This has been done for the SRS by several researchers, such as Duku et al. (2013), who developed a 30-item SRS form, by (Gau et al., 2013), who examined a four-factor 60-item SRS form, and by Moul et al. (2015) and Sturm et al. (2017), who both developed a 16-item SRS form using different items. Because Sturm et al., (2017) developed a short SRS form using item response theory, which was found to discriminate adequately between children without and with ASD in another study (Nguyen et al., 2019), we decided to use this model in the current study. Although the AIC indicated the best model fit for the SRS short version, compared to the models using the long version of the SRS, the remaining model fit indices were rather poor. Taking into account the widespread use of the SRS and the findings of the current study in terms of Cronbach's alphas, the correlations of the SRS with other instruments, and especially the results of the factor analyses, that are in line with previous research, further scale refinement is indicated.

4.1. Limitations

There are several limitations that must be considered when reading and interpreting the findings of the current study. First, we did not examine the criterion validity of the SRS scores with ASD diagnoses, as those findings would have been somewhat misleading. The sample included participants with different diagnoses, such as ASD, ADHD, and ID, without co-occurring diagnoses, as well as participants with epilepsy and other diagnoses and comorbidities. Typically developing participants were not included. In addition, ASD diagnoses were based on ICD-10 criteria obtained from medical records. It would have been beneficial to have the SRS compared directly with specific ASD measures, such as ADOS and ADI-R. However, only scant information was available regarding those scores in the current sample. Furthermore, the sample that consisted of participants with different diagnoses and with perhaps varying autistic traits might also have contributed to the somewhat poor model fit in the different CFAs. However, previously conducted research has also used samples with mixed diagnoses to examine SRS (e.g., Moul et al., 2015; Wang et al., 2012). The SRS is a measure of deficits in social behavior and is relevant not only for the assessment of patients with ASD but also for patients with other disorders, such as epilepsy or ID (Aburahma et al., 2021). Aburahma et al. (2021) proposed, for example, that the SRS should be used 'as a social impairment screening tool to identify children requiring referral for comprehensive psychiatric assessment' (p. 1171). Another consideration that must be taken into account when interpreting the findings is that the sample size was relatively small in relation to the complex models that were analyzed using CFA (Kyriazos, 2018). However, the CFA analyses were probably adequately powered taking into account the sample size, the number of indicators, and the few missing datapoints, and thus not the reason for the poor model fit (Wolf et al., 2013). It would have been desirable to use the weighted least square mean and variance adjusted (WLSMV) estimator; however, there were model conversion problems for the five-factor model of the original SRS, which might be related to the relatively small sample size. Using an ML estimator might have led to less accurate factor loadings (Li, 2016), standard errors (DiStefano, 2002), and chi-square test results (Beauducel & Herzberg, 2006).

4.2. Future research

Future research should continue to examine the psychometric properties of the SRS in larger samples and aim at improving its factor structure. A promising alternative to a 65-item full version of the SRS could be a shortened version. In particular, the psychometric properties of a reduced version should be examined in the future, as it might be a useful screening instrument in clinical practice as well as for research.

4.3. Conclusion

The SRS is a widely used quantitative measure of autistic traits. Studies of the scale's psychometric properties are important to shed light on the strengths and weaknesses of the instrument and provide input for further developmental needs in this field. The majority of the SRS subscales in the current study had adequate internal consistency. The findings from the present study in a mixed neurodevelopmental/neurological sample found the SRS parent version to be reliable overall and show meaningful overlap with related measures. However, the lack of SRS construct validity in terms of underling factor structures (one-, five- and second-order five-factor models) on the 65-item long form, indicates that the SRS would profit from further scale refinement. Although the short version had poor model fit in the current study, generally reducing questions from a 65-item to a shortened questionnaire by applying scientific methods that lead to an equally sound or even improved instrument could be useful and save time both in clinical and research praxis. This is especially important for the SRS, which has been found to have poor model fit in its original form in several studies. Future research is needed that examines the psychometric properties of an improved version of the SRS.

CRediT authorship contribution statement

Kjersti Åby Bergquist conducted the analyses and wrote a first draft. Sabine Kaiser conducted the analyses, revised the manuscript critically, and wrote the final version of the article. Marianne Berg Halvorsen designed the study, conducted the data acquisition, and revised the manuscript critically. All authors approved the final version of the manuscript.

Declarations of interest

None. There are no actual or potential conflicts of interest regarding this study.

Data availability

The data underlying this paper cannot be shared because they contain patient information.

Acknowledgments

The authors would like to thank the participants who contributed to the dataset.

Funding

This study was supported by the Northern Norway Regional Health Authority (grant number PFP1178–14). The study was also supported in part by The Norwegian Directorate of Health and the Regional Centre for Child and Youth Mental Health and Child Welfare North at UiT The Arctic University of Norway, Tromsø, Norway.

Informed consent

Written informed consent was obtained from all participants before inclusion in the study. All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

References

Aburahma, S. K., Hammouri, H., Hazaimeh, E., Jbarah, O., Nassar, A., Almasri, A., … Bashtawi, M. (2021). Social impairment in children with epilepsy assessed by the social responsiveness scale. *Clinical Child Psychology and Psychiatry, 26*(4), 1170–1181. https://doi.org/10.1177/13591045211033176

Aman, M. G., & Singh, N. (1994). *Aberrant Behavior Checklist. Community (abc).* Slosson Educational Publications, Inc,.

American Psychiatric Association, 2013.

Beauducel, A., & Herzberg, P. Y. (2006). On the performance of maximum likelihood versus means and variance adjusted weighted least squares estimation in cfa. *Structural Equation Modeling: A Multidisciplinary Journal, 13*(2), 186–203. https://doi.org/10.1207/s15328007sem1302_2

Bolte, S., Poustka, F., & Constantino, J. N. (2008). Assessing autistic traits: Cross-cultural validation of the social responsiveness scale (srs). *Autism Research, 1*(6), 354–363. https://doi.org/10.1002/aur.49

Cen, C. Q., Liang, Y. Y., Chen, Q. R., Chen, K. Y., Deng, H. Z., Chen, B. Y., & Zou, X. B. (2017). Investigating the validation of the chinese mandarin version of the social responsiveness scale in a mainland china child population. *BMC Psychiatry, 17*(1), 51. https://doi.org/10.1186/s12888-016-1185-y

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. https://doi.org/10.1192/bjp.bp.107.040196

Cheon, K. A., Park, J. I., Koh, Y. J., Song, J., Hong, H. J., Kim, Y. K., … Kim, Y. S. (2016). The social responsiveness scale in relation to dsm iv and dsm5 asd in korean children. *Autism Research, 9*(9), 970–980. https://doi.org/10.1002/aur.1671

Constantino, J. N., & Gruber, C. P. (2005). *The Social Responsiveness Scale Manual.* Western Psychological Services.

Constantino, J. N., & Gruber, C. P. (2012). *Social Responsiveness Scale: Srs-2.* Western Psychological Services.

DiStefano, C. (2002). The impact of categorization with confirmatory factor analysis. *Structural Equation Modeling: A Multidisciplinary Journal, 9*(3), 327–346. https://doi.org/10.1207/S15328007SEM0903_2

DiStefano, C., & Hess, B. (2005). Using confirmatory factor analysis for construct validation: An empirical review. *Journal of Psychoeducational Assessment, 23*(3), 225–241. https://doi.org/10.1177/073428290502300303

Duku, E., Vaillancourt, T., Szatmari, P., Georgiades, S., Zwaigenbaum, L., Smith, I. M., … Pathways in A. S. D. Study Team. (2013). Investigating the measurement properties of the social responsiveness scale in preschool children with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 43*(4), 860–868. https://doi.org/10.1007/s10803-012-1627-4

European Federation of Psychologists' Associations, 2013, Efpa review model for the description and evaluation of psychological and educational tests: Test review form and notes for reviewers (version 4.2.6). EFPA.

Fombonne, E., Marcin, C., Bruno, R., Tinoco, C. M., & Marquez, C. D. (2012). Screening for autism in mexico. *Autism Research, 5*(3), 180–189. https://doi.org/10.1002/aur.1235

Gau, S. S. F., Liu, L. T., Wu, Y. Y., Chiu, Y. N., & Tsai, W. C. (2013). Psychometric properties of the chinese version of the social responsiveness scale. *Research in Autism Spectrum Disorders, 7*(2), 349–360. https://doi.org/10.1016/j.rasd.2012.10.004

Goodman, R. (1997). The strengths and difficulties questionnaire: A research note. *Journal of Child Psychology and Psychiatry, 38*(5), 581–586. https://doi.org/10.1111/j.1469-7610.1997.tb01545.x

Halvorsen, M., Aman, M. G., Mathiassen, B., Brøndbo, P. H., Steinsvik, O. O., & Martinussen, M. (2019). Psychometric properties of the norwegian aberrant behavior checklist and diagnostic relationships in a neuro-pediatric sample. *Journal of Mental Health Research in Intellectual Disabilities, 12*(3–4), 234–255. https://doi.org/10.1080/19315864.2019.1630872

Hooper, D., Coughlan, J., & Mullen, M. (2008). Structural equation modelling: Guidelines for determining model fit structural equation modelling: Guidelines for determining model fit. *Electronic Journal on Business Research Methods, 6*, 53–60.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Hus, V., Bishop, S., Gotham, K., Huerta, M., & Lord, C. (2013). Factors influencing scores on the social responsiveness scale. *Journal of Child Psychology and Psychiatry, 54*(2), 216–224. https://doi.org/10.1111/j.1469-7610.2012.02589.x

Kaiser, S., & Halvorsen, M. B. (2022). The strengths and difficulties questionnaire self-report-, parent-, and teacher version in children with intellectual and developmental disabilities. *Research in Developmental Disabilities, 123*, Article 104194. https://doi.org/10.1016/j.ridd.2022.104194

Kyriazos, T. A. (2018). Applied psychometrics: Sample size and sample power considerations in factor analysis (efa, cfa) and sem in general. *Psychology, 09*, 2207–2230.

Li, C.-H. (2016). Confirmatory factor analysis with ordinal data: Comparing robust maximum likelihood and diagonally weighted least squares. *Behavior Research Methods, 48*(3), 936–949. https://doi.org/10.3758/s13428-015-0619-7

Lord, C., Risi, S., Lambrecht, L., Cook, E. H., Leventhal, B. L., DiLavore, P. C., … Rutter, M. (2000). The autism diagnostic observation schedule—generic: A standard measure of social and communication deficits associated with the spectrum of autism. *Journal of Autism and Developmental Disorders, 30*(3), 205–223. https://doi.org/10.1023/A:1005592401947

Lord, C., Rutter, M., DiLavore, P., Risi, S., Gotham, K., & Bishop, S. (2012). *Autism Diagnostic Observation Schedule–2nd Edition (ados-2)*. Western Psychological Corporation,.

Matsunaga, M. (2010). How to factor-analyze your data right: Do's, don'ts, and how-to's. *International Journal of Psychological Research, 3*(1), 97–110. https://doi.org/10.21500/20112084.854

Moul, C., Cauchi, A., Hawes, D. J., Brennan, J., & Dadds, M. R. (2015). Differentiating autism spectrum disorder and overlapping psychopathology with a brief version of the social responsiveness scale. *Child Psychiatry and Human Development, 46*(1), 108–117. https://doi.org/10.1007/s10578-014-0456-4

Nguyen, P. H., Ocansey, M. E., Miller, M., Le, D. T. K., Schmidt, R. J., & Prado, E. L. (2019). The reliability and validity of the social responsiveness scale to measure autism symptomology in vietnamese children. *Autism Research, 12*(11), 1706–1718. https://doi.org/10.1002/aur.2179

Pine, E., Luby, J., Abbacchi, A., & Constantino, J. N. (2006). Quantitative assessment of autistic symptomatology in preschoolers. *Autism, 10*(4), 344–352. https://doi.org/10.1177/1362361306064434

Raven, J. (2004). *The Raven's – Educational Coloured Progressive Matrices and Crichton Vocabulary Scale. Norwegian manual supplement.* Pearson Assessment,.

Rutter, M., Lecouteur, A., & Lord, C. (2003). *The Autism Diagnostic Interview - Revised (adi-r)*. Western Psychological Services,.

Shahrivar, Z., Tehrani-Doost, M., Davoodi, E., Hosseiniani, T., Tarighatnia, H., Momen, S., … Hajirezaei, H. (2020). The reliability of the social responsiveness scale-2 in an iranian typically developing group of children. *Iranian Journal of Psychiatry, 15*(1), 41–46. https://doi.org/10.18502/ijps.v15i1.2438

Sparrow, S. S., Cicchetti, D. V., & Balla, D. A. (2011). *Vineland Adaptive Behavior Scales (2nd ed.). Norwegian Manual Supplement.* Pearson Assessment,.

Sturm, A., Kuhfeld, M., Kasari, C., & McCracken, J. T. (2017). Development and validation of an item response theory-based social responsiveness scale short form. *Journal of Child Psychology and Psychiatry, 58*(9), 1053–1061. https://doi.org/10.1111/jc

Surén, P., Havdahl, A., Øyen, A.-S., Schjolberg, S., Reichborn-Kjennerud, T., Magnus, P., … Stoltenberg, C. (2019). Diagnostisering av autismespekterforstyrrelser hos barn i norge. *Tidsskrift for den Norske Laegeforening, 139*(14). https://doi.org/10.4045/tidsskr.18.0960

Tavakol, M., & Dennick, R. (2011). Making sense of cronbach's alpha. *International Journal of Medical Education, 2*, 53–55. https://doi.org/10.5116/ijme.4dfb.8dfd

Wang, J., Lee, L. C., Chen, Y. S., & Hsu, J. W. (2012). Assessing autistic traits in a taiwan preschool population: Cross-cultural validation of the social responsiveness scale (srs). *Journal of Autism and Developmental Disorders, 42*(11), 2450–2459. https://doi.org/10.1007/s10803-012-1499-7

Wechsler, D. (2007). *Wechsler Abbreviated Scale of Intelligence (wasi). Norwegian Manual Supplement*. Pearson Assessment.

Wechsler, D. (2008aaa). *Wechsler Adult Intelligence Scale* (4th ed..,). Pearson Assessment.

Wechsler, D. (2008bbb). *Wechsler Preschool and Primary Scale of Intelligence (wppsi-iii)* (3rd ed..,). Pearson Assessment.

Wechsler, D. (2009). *Wechsler Intelligence Scale for Children (wisc-iv)* (4th ed..,). Pearson Assessment.

Wechsler, D. (2012). *Wechsler Preschool and Primary Scale of Intelligence (wppsi-iv)* (4th ed..,). Pearson Assessment.

Wigham, S., McConachie, H., Tandos, J., & Le Couteur, A. S. (2012). The reliability and validity of the social responsiveness scale in a uk general child population. *Research in Developmental Disabilities, 33*(3), 944–950. https://doi.org/10.1016/j.ridd.2011.12.017

Wolf, E. J., Harrington, K. M., Clark, S. L., & Miller, M. W. (2013). Sample size requirements for structural equation models: an evaluation of power, bias, and solution propriety. *Educational and Psychological Measurement, 76*(6), 913–934. https://doi.org/10.1177/0013164413495237

World Health Organization. (1992). *International Statistical Classification of Diseases and Related Health Problems (icd-10)* (10th ed..,). World Health Organization.

World Health Organization. (2018). *International Statistical Classification of Diseases and Related Health Problems: 11th Revision (icd-11)*. World Health Organization. ⟨https://icd.who.int/browse11/l-m/en⟩.