J Autism Dev Disord (2013) 43:860–868
DOI 10.1007/s10803-012-1627-4

Investigating the Measurement Properties of the Social Responsiveness Scale in Preschool Children with Autism Spectrum Disorders

Eric Duku · Tracy Vaillancourt · Peter Szatmari · Stelios Georgiades · Lonnie Zwaigenbaum · Isabel M. Smith · Susan Bryson · Eric Fombonne · Pat Mirenda · Wendy Roberts · Joanne Volden · Charlotte Waddell · Ann Thompson · Teresa Bennett · the Pathways in ASD Study Team

Published online: 23 August 2012

Abstract The purpose of this study was to examine the measurement properties of the Social Responsiveness Scale in an accelerated longitudinal sample of 4-year-old preschool children with the complementary approaches of categorical confirmatory factor analysis and Rasch analysis. Measurement models based on the literature and other hypothesized measurement models which were tested using categorical confirmatory factor analysis did not fit well and were not unidimensional. Rasch analyses showed that a 30-item subset met criteria of unidimensionality and invariance across item, person, and over time; and this subset exhibited convergent validity with other child outcomes. This subset was shown to have enhanced psychometric properties and could be used in measuring social responsiveness among preschool age children with Autism Spectrum Disorders.

Keywords Social Responsiveness Scale · Autism spectrum disorders · Measurement · Confirmatory factor analysis · Rasch analyses · Structural equation modelling

E. Duku · P. Szatmari · S. Georgiades · A. Thompson · T. Bennett
Offord Centre for Child Studies, McMaster University, Hamilton, ON, Canada

T. Vaillancourt (✉)
Faculty of Education and School of Psychology, University of Ottawa, 145, Jean-Jacques-Lussier Private, Ottawa, ON K1N 6N5, Canada
e-mail: tracy.vaillancourt@uottawa.ca

L. Zwaigenbaum · J. Volden
University of Alberta, Edmonton, AB, Canada

I. M. Smith · S. Bryson
Dalhousie University/IWK Health Centre, Halifax, NS, Canada

E. Fombonne
Montreal Children's Hospital, Montreal, QC, Canada

P. Mirenda
University of British Columbia, Vancouver, BC, Canada

W. Roberts
University of Toronto, Toronto, ON, Canada

C. Waddell
Simon Fraser University, Vancouver, BC, Canada

Introduction

Autism spectrum disorders (ASD) are neurodevelopmental conditions with considerable morbidity and costs to individuals, their families, and society (Charman 2007). ASD affects roughly 1 in 88 preschool aged children (Centers for Disease Control and Prevention 2012). Children are diagnosed with ASD based on impairments in social interaction and communication, as well as a pattern of repetitive or stereotypic behaviour and interests (APA 2000). The diagnosis of ASD usually involves the use of direct observation instruments (e.g., Autism Diagnostic Observation Schedule, ADOS; Lord et al. 2000), parent interviews (e.g., Autism Diagnostic Interview-Revised, ADI-R; Lord et al. 1994), and additional information from independent sources such as teachers, as well as clinical judgment (Newschaffer et al. 2007). Questionnaires such as the Social Responsiveness Scale (SRS; Constantino and Gruber 2005) are frequently used to obtain supplementary information about the child's symptoms.

The SRS is a 65-item quantitative scale that measures the severity of social impairment symptoms related to ASD in individuals between 4 and 18 years of age. In addition to a total severity score, the SRS has five conceptually

derived subscales: Social Awareness, Social Cognition, Social Communication, Social Motivation, and Autistic Mannerisms (Constantino and Gruber 2005). The SRS has been shown to have good sensitivity and specificity and to be informative for differential diagnosis, successfully distinguishing ASD from other childhood psychiatric conditions (Constantino and Gruber 2005). The SRS can be completed in 20 minutes by parents, teachers, or childcare providers who have observed the child's interactions with peers in naturalistic settings. In contrast to other measures of ASD symptoms (e.g., ADOS and ADI-R), the measurement framework for the SRS models autism as a unidimensional phenotype rather than a multidimensional and/or categorical construct (Constantino and Gruber 2005).

Constantino and colleagues have examined the psychometric properties and utility of the SRS across different samples of children varying in age from 4 to 18 years. For example, Constantino et al. (2000) examined the discriminant validity of the SRS in 158 child psychiatric patients, with and without ASD, and a control sample of 287 children randomly selected from a school district. The factor structure of the SRS was also examined using Latent Class Analysis. Results indicated that a one-factor solution, which explained 70 % of the variance, best fit the data. Constantino et al. (2004) re-examined the factor structure using principal component analysis with a clinical sample of 168 children and 259 administrations of the SRS (parent and teacher) and once again showed that a one-factor solution, explaining about 35 % of the variance, best fit the data. Based on these and other analyses, Constantino et al. (2004) concluded that there was no evidence of separate independent subdomains of impairment associated with autism as measured by the SRS in 4- to 18-year-olds. This finding is inconsistent with those reported in previous studies that have described the ASD symptom phenotype as multidimensional, comprising the three domains of social deficits, communication deficits, and fixated interests/repetitive behaviour (Frazier et al. 2008; Georgiades et al. 2007). It should also be noted that the difference in amounts of variance explained could be attributed to the two methodological approaches used—factor analysis (FA) and latent class analysis (LCA).

Although the SRS is commonly used to assess severity of autistic social impairment symptoms in children and youth with ASD, most of the psychometric work has been completed in general population samples. More research has been needed to address the measurement properties of the SRS in children diagnosed with ASD. Specifically, data are needed on the measurement model of the SRS (i.e., uni- vs. multi-dimensionality), and on the longitudinal invariance and psychometric stability of the SRS. It would also be informative to examine the measurement properties of the SRS in younger children, given that ASD is being identified more often in the early preschool years.

Earlier work on the psychometric properties of the SRS used classical test theory approaches such as linear factor analyses, correlations, and item-total correlations (or internal consistency). Although these tests are informative at the scale level, they do not allow the examination of response patterns for individual items. As a complementary approach, Rasch analysis using a latent trait model (with single items treated as indicators) is increasingly being used in health sciences research because it allows researchers to test for response patterns for individual items, for individual-person estimates, and for individual item and person fits and residuals (Hagquist et al. 2009). Rasch's unidimensional measurement model reflects a fundamental feature of measurement; an instrument should work the same way for all individuals (Andrich 1988). Rasch analysis is usually employed in developing and examining measurement instruments, and is useful in analysing the psychometric properties of composite measures that are considered to capture unidimensional constructs, such as the SRS. It is appropriate for Likert-like response items as in the SRS and can be classified as an Item Response Theory (IRT) model. Rasch analysis involves testing whether patterns of responses to items conform to model expectations. To our knowledge, Rasch analysis has not been used to evaluate the measurement properties of instruments assessing ASD symptoms.

The objective of this study was to use multiple methods to investigate the measurement properties of the 4–18-year version of the SRS in a sample of newly diagnosed 4-year-old preschool children with ASD. This objective was achieved by examining: (1) competing measurement models of the SRS in a clinical ASD sample; (2a) an empirical measurement model of the underlying structure of the SRS using Rasch analysis; (2b) the longitudinal invariance of the resulting measurement model of the SRS; and (3) the convergent and discriminant validity of the resulting measurement model with concurrent child outcome measures.

Methods

Participants

Participants were recruited into an on-going longitudinal study of children with ASD (*Pathways in ASD* study) through regional ASD referral centres across Canada (Halifax, Montreal, Hamilton, Edmonton, and the Greater Vancouver/Fraser Valley regions of British Columbia). The study was approved by the Research Ethics Boards at all sites. Families willing to participate provided informed

consent prior to joining the study. Participants included 339 children younger than 60 months (mean age at consent = 39.8 months, SD = 8.9) with a recent diagnosis of ASD. Inclusion criteria for participation in the Pathways Study were as follows: (a) recent (i.e., within 4 months) clinical diagnosis of ASD, confirmed by both the Autism Diagnostic Observation Schedule (ADOS; Lord et al. 2000) and the Autism Diagnostic Interview-Revised (ADI-R; Risi et al. 2006), as well as a diagnosis assigned by a clinician using DSM-IV criteria (APA 2000); and (b) chronological age equal to, or older than, 2 years and equal to, or younger than, 5 years and 0 months. Children were excluded from the study if any of the following conditions were present: (a) cerebral palsy or other neuromotor disorders interfering with study assessments; (b) any known genetic or chromosomal abnormality; or (c) severe visual or hearing impairment. To ensure independence of observations, only one child per family was recruited to the study.

There were both cross-sectional and longitudinal components to this study. For the cross-sectional component, this paper focused on analyses involving an accelerated longitudinal sample (n = 205; 177 boys and 28 girls) of 4-year-olds derived by combining data from baseline (T1; n = 61), 6-month follow-up (T2; n = 75) and one-year follow-up (T3; n = 69), drawn from the whole *Pathways in ASD* cohort. The accelerated longitudinal sample was used so that data could be combined from children of the same age (4 years), from the three "age at time of diagnosis" cohorts at different assessment points (4-year-olds at T1, T2 and T3; see Fig. 1). Each child provided data at only one time point of assessment. For the longitudinal component, the accelerated longitudinal data (T1, T2, and T3) were used as the baseline data along with follow-up data at 6 years (T4). Parents of participants provided ratings for some instruments while other ratings were based on observer reports. Most (93.3 %) parent reports were obtained from mothers, with a mean maternal age at consent of 35.3 years (SD = 5.3).

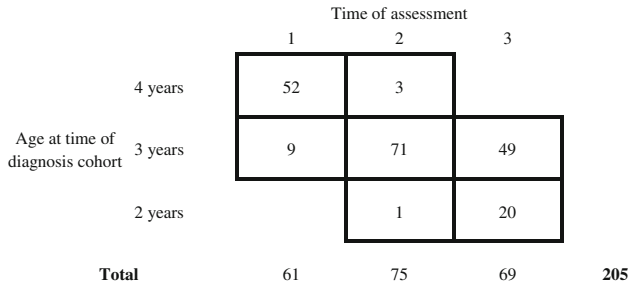

Fig. 1 Accelerated longitudinal design used for children 4 years old at time of assessment

Measures

Social Responsiveness Scale (SRS)

The SRS provides a picture of a child's atypical social behaviour including social awareness, social information processing, reciprocal communication, social anxiety or avoidance, and autistic preoccupations and traits (Constantino and Gruber 2005). Ratings on the items are provided by the child's caregiver on a scale from 1 (not true) to 4 (almost always true) based on the frequency (not the intensity) of the behaviour. The items vary in degree of abnormality since some inquire about mildly abnormal behaviours whereas others inquire about the severely abnormal. Higher total scores indicate greater severity of social impairment.

Child Behavior Checklist 1.5–5 (CBCL1.5-5)

The 99-item CBCL is a widely-used norm-referenced instrument that can evaluate a wide range of internalizing and externalizing disorders, based on six subscales (Emotionally reactive, Anxious/depressed, Somatic complaints, Withdrawn, Attention problems, Aggressive behavior; Achenbach and Rescorla 2000). The CBCL is completed by parents or teachers based on observations of the child's behaviour in the previous 2 months. Scale and subscale scores are summed and converted to T-scores. The CBCL has good test–retest and inter-rater reliability for all scales and subscales. The authors also report evidence of discriminative, convergent, and predictive validities (Achenbach and Rescorla 2000).

Repetitive Behavior Scale-Revised (RBS-R)

The RBS-R is a clinical rating scale that measures the presence and severity of a range of restricted, repetitive behaviours that are associated with ASD (Bodfish et al. 2000). It is completed by parents and provides a quantitative, continuous measure of repetitive behaviours. It consists of 43 items distributed across six conceptually derived subscales: Stereotyped behaviour, Self-injurious behaviour, Compulsive behaviour, Routine behaviour, Sameness behaviour, and Restricted behaviour. Mirenda et al. (2010) validated the utility of the RBS-R as a measure of repetitive behaviours in this sample of preschool children with ASD.

Preschool Language Scale: Fourth Edition (PLS-4)

The PLS-4 is a norm-referenced and comprehensive language test for identifying children with a language disorder or delay. It is administered individually to children between birth and age 6 years and 11 months, or to older children who function developmentally within this age

range (Zimmerman et al. 2002). The PLS-4 was used to obtain an index of early syntax and semantic skill in this sample of preschool children with ASD (Volden et al. 2011).

Vineland Adaptive Behavior Scale Second Edition (VABS-II)

The VABS-II was designed to assess functioning from birth to 18 years in the domains of Communication, Daily living skills, Socialization, and Motor skills (Sparrow et al. 2005). Scores from domains and sub-domains permit the comparison of specific profiles of adaptive behaviours in these groups. The VABS-II is administered to parents or caregivers using a semi-structured interview format. Open-ended questions are used to gather detailed information and promote rapport between interviewers and respondents. The VABS-II has been shown to have adequate reliability and validity (Sparrow et al. 2005).

Analyses

The analyses were conducted in three stages. First, the measurement properties of the SRS were examined by considering internal consistencies of the total SRS score and subscales in the sample. Second, the hypothesized unidimensional measurement model and other hypothesized measurement models were evaluated using categorical confirmatory factor analysis (CCFA) in a structural equation modelling (SEM) framework with Mplus version 5.1 (Muthén and Muthén 2008a, b). Tests of goodness-of-fit of the models were evaluated using the criteria described by Hu and Bentler (1999), who recommend conducting several goodness-of-fit tests and reporting their resulting indices. Third, based on the results of the CCFAs, the measurement model of the SRS was examined using Rasch analysis of the 65 items as a complementary approach.

The Rasch model for Likert-like items with ordered categories, called the polytomous Rasch model, is also known as the rating scale model (Andrich 1988). The rating scale model assumes equal intervals between adjacent categories across all items whereas the partial credit model does not impose any restrictions on the intervals (Andrich 1988). The Rasch polytomous model is suitable for analysis of the SRS which has items with ordered Likert-like categories.

The Rasch unidimensional measurement model assumes that the probability that a particular individual will endorse an item is a logistic function of the relative distance between the location of the item and the person location. In other words, the probability that a parent or caregiver will endorse an item is a logistic function of the difference between the child's level or severity of autistic symptoms and the level of severity of autistic symptoms expressed by the item. The response patterns are tested against the expected pattern and a variety of fit statistics are used to determine how well the responses fit the pattern (Hagquist et al. 2009). A good fit of the response pattern means that for the same latent trait, the probability of endorsing a more severe item is higher than the probability of endorsing a less severe item. Rasch's work has been extensively reviewed by other statisticians and methodologists and is well suited to the examination of the measurement properties of instruments such as the SRS (Andrich 2004; Hagquist et al. 2009; Tennant and Conaghan 2007; Wright 1977).

Using the RUMM2020 software (Andrich et al. 2007), the Rasch analysis of the SRS was iterative and included: (a) a test of which polytomous version (rating scale or partial credit model) was appropriate using the Likelihood Ratio Test between models; (b) an overall test of how well the SRS data fit the Rasch model; (c) stepwise deletion of items that showed local dependency based on correlations of residuals over 0.3 after removal of the Rasch model; (d) stepwise deletion of poor-fitting items with extreme item-fit residuals (over $\pm$ 2.5); (e) deletion of items with disordered thresholds (those thresholds between response options of items that do not display an increasing level of the trait); (f) deletion of cases with extreme person-fit residuals (over $\pm$ 2.5), and testing for differential item response patterns or differential item functioning (DIF) across the three data points of the accelerated longitudinal sample; and finally, (g) a test of how well the remaining items and cases fit the Rasch model and the assumptions of invariance and unidimensionality.

Longitudinal invariance was examined with Rasch analysis using the accelerated longitudinal data (T1, T2, and T3) as the baseline and follow-up data at 6 years (T4). The Rasch model was considered to be an adequate fit if the summary and individual χ^2 statistics were non-significant ($p > 0.05$) after adjusting for multiple testing using the Bonferroni correction (Hagquist et al. 2009). Evidence of differential item functioning (DIF) was assessed by analysing the residuals with the three data points and the estimated latent score as covariates (Hagquist and Andrich 2004).

Finally, convergent validity was examined by comparing the resulting total score from the retained items of the SRS to the 65-item total score and to concurrent child outcome measures (the CBCL, VABS-II, RBS-R and PLS-4).

Results

Examining Measurement Models of the SRS

Internal consistency of the 65-item SRS total raw score was good (Cronbach's $\alpha = 0.93$), indicating strong

Table 1 Internal consistencies of the data for the total and subscales raw scores of the SRS in 4-year-olds (n = 205)

Scale/subscale	# Items	Cronbach's alpha
Total	65	0.93
Social awareness	8	0.60
Social cognition	12	0.72
Social communication	22	0.85
Social motivation	11	0.70
Autistic mannerisms	12	0.79

intercorrelations between items and the total raw score. The internal consistencies for the SRS subscales for this sample were also acceptable (above 0.70) except for the Social Awareness subscale, for which the internal consistency was 0.60 (see Table 1).

The goodness-of-fit indices from the CCFA indicated a poor fit for the one-factor (unidimensional) structure comprising all 65 items of the SRS (χ^2 (113) = 431.975, CFI = 0.686, TLI = 0.761, RMSEA = 0.119). Two other tested models, a 5-factor first-order model and a 5-factor second-order model with the subscales as factors, also did not fit well (see Table 2). Single-order unidimensional measurement models were also tested for the SRS subscales. As summarized in Table 2, none of these models met criteria for adequate fit suggested by Hu and Bentler (1999): RMSEA value less than or equal to 0.6, and/or CFI (or TLI) greater than or equal to 0.9, and/or a Chi-square statistic with p value greater than 0.05.

Rasch Analysis: Measurement Model

Since none of the hypothesized measurement models of the SRS provided a good fit to the data, Rasch analysis was used as a complementary approach to examine the measurement properties of the 65-item SRS, including assessing dimensionality, response patterns for individual items, individual person estimates, and individual item and person residuals with fit indices. Results indicated that the partial credit model was appropriate for the data based on the Likelihood Ratio Test ($\chi^2(127) = 434.2$, $p < 0.001$). The person separation index (PSI; a measure of reliability, similar to Cronbach's α) of 0.93 indicated high internal consistency. However, the overall fit of the model evaluated using the latent-trait Chi-square statistic from Table 3 for the 65-item SRS was poor ($\chi^2(130) = 130$, $p < 0.001$). A test of local dependencies (or intercorrelations) between items based on a residual PCA (after removing the Rasch model) indicated that 10 items had correlations over 0.3 with other items. Using an iterative process, 9 items were initially excluded from the item set because they did not fit the Rasch model (i.e., item-fit residuals greater than ±2.5) or were locally dependent on other items (i.e., residual correlations greater than 0.3). In the second phase, 11 of the remaining 46 items had disordered thresholds and 5 other items did not fit the Rasch model, so an additional 16 items were excluded from the item set, for a total of 35 items. For example, the item "doesn't recognize when others are trying to take advantage of him/her" displayed a fit to the Rasch model whereas another item ("knows when he/she is too close to someone or is invading someone's space") displayed under-discrimination to the Rasch model. Of the 35 items excluded, 5 were excluded from the Social Awareness subscale, 6 from the Social Cognition subscale, 15 from the Social Communication subscale, 5 from the Social Motivation subscale and 4 from the Autistic Mannerisms subscale.

In the final phase of the iterative procedure, 24 children were also excluded from the analyses because their person-fit residuals were over ± 2.5 (i.e., their expected person estimates were 2.5 standardized units away from the observed person estimates). The 30-item subset comprised 8 items which were retained from the Autistic Mannerisms subscale, 7 from the Social Communication subscale, 6 from each of the Social Motivation and Social Cognition subscales, and 3 from the Social Awareness subscale (see "Appendix"). The statistics for the resulting model with 30 items are presented in Table 3. As shown, the item-trait test was non-significant ($\chi^2(60) = 76.008$, $p = 0.08$), indicating that the data fit the Rasch model and that assumptions of invariance (item and person) and unidimensionality held.

Table 2 Goodness-of-fit statistics for models tested using categorical confirmatory factor analyses with four-year-olds from T1, T2, and T3 (n = 205)

	Chi-sq, *df*, *p* value	CFI	TLI	RMSEA
1-Factor model	431.975, 113, <0.001	0.686	0.761	0.119
5-Factor model	419.516, 113, <0.001	0.698	0.770	0.116
2nd Order 5-factor model	420.940, 113, <0.001	0.697	0.769	0.116
Social awareness	47.515, 15, <0.001	0.767	0.720	0.104
Social cognition	141.071, 28, <0.001	0.654	0.666	0.142
Social communication	264.904, 65, <0.001	0.675	0.823	0.124
Social motivation	162.396, 26, <0.001	0.680	0.693	0.162
Autistic mannerisms	120.732, 32, <0.001	0.833	0.875	0.117

CFI Comparative Fit index, *TLI* Tucker-Lewis Index, *RMSEA* root mean square error of approximation, *SRMR* standardized root mean square residual

Table 3 Summary of test of fit statistic for Rasch analysis of the SRS using selected sample of 205 4-year-olds

	N	Item residual value (SD)	Person residual value (SD)	Item-trait Chi-square	Degrees of freedom	*p* value	Person Separation Index	Power of fit test
65-Item partial credit model	205	0.000 (0.626)	−0.333 (0.577)	248.8	130	<0.001	0.931	Excellent
65-Item rating scale model	205	0.000 (0.763)	−0.381 (0.579)	238.767	130	<0.001	0.931	Excellent
30-Item partial credit model	181	0.000 (0.583)	−0.387 (0.667)	76.008	60	0.080	0.884	Excellent
30-Item rating scale model	181	0.000 (0.636)	−0.391 (0.667)	81.256	60	0.035	0.884	Excellent

Rasch Analysis: Invariance (Across Groups and Over Time)

Examination of DIF for the 3 data points from which 4-year-old participants were drawn (i.e., baseline, 6 and 12 months later) showed no evidence of uniform or non-uniform DIF since no comparisons using analyses of variance were statistically significant. This indicated invariance of the measurement model across the three data points in the accelerated longitudinal sample.

Further examination of DIF using data from the accelerated longitudinal sample (i.e., at age 4) as the baseline and data from time 4 (i.e., 12–36 months later) as follow-up showed that other than uniform DIF for "doesn't recognize when others are trying to take advantage of him/her", there was no evidence of uniform or non-uniform DIF, indicating invariance across time. No comparisons using analyses of variance with Bonferroni correction for multiple comparisons were statistically significant.

Convergent Validity

Convergent validity was investigated by examining the correlations between the 30-item set with concurrent child outcome measures in the study for the children in the accelerated longitudinal sample. The strength of the association between the 30-item subset total score and the 65-item SRS total score ($r = 0.94$, $p < 0.001$) indicated that the 30-item subset total score accounted for approximately 88 % of the variance of the 65-item SRS total score. The 30-item subset total score was positively correlated with the CBCL and RBS-R (r from 0.65 to 0.67, see Table 4) and negatively related with the VABS-II ($r = -0.33$, $p < 0.001$). The PLS-4 had no significant relationship with the 30-item subset total scores ($r = -0.09$, $p = 0.191$).

Table 4 Correlations between total score of the 30-item subset with concurrent child outcome measures for the accelerated longitudinal sample of 4-year-olds

Child outcome measures	30-item subset total raw score
SRS (65-item) Total raw score	0.94**
CBCL internalizing problems: total	0.68**
CBCL externalizing problems: total	0.65**
RBS-R overall mean score	0.67**
VABS-II adaptive behaviour composite standard score	−0.33**
PLS-4 total language standard score	−0.09

** $p < 0.01$ level (2-tailed)

Discussion

This is the first comprehensive study to assess the measurement properties of the SRS in a clinical sample of recently diagnosed 4-year-old preschool children with ASD. It is also the first study to use Rasch modelling to examine the properties of the SRS in an ASD sample. Examination of the measurement properties of the SRS in 4-year-olds is important given that many children are being referred for ASD assessment by this age (Chawarska et al. 2007) and that the SRS may be an informative data source for clinicians. However, poor fit statistics and indices from the hypothesized unidimensional CCFAs showed that the 65-item SRS could not be characterized as unidimensional in our study. The implication of this finding is that one cannot assume measurement equivalence for any measure to be used with ASD children across as wide an age span as the SRS suggests.

Examination of the SRS data using the complementary approach of Rasch analysis also confirmed that the 65-item SRS could not be characterized as unidimensional and that the items did not form a well-fitting measurement model. It is possible that the lack of unidimensionality of the 65-item SRS arose because the covariance between the items could not be explained by a single underlying construct. It is also possible that item properties differed according to some grouping variable or item redundancy or dependency. Another possible reason for the poor fit of the measurement models may be that certain SRS items are less relevant for younger children, e.g., "has good personal hygiene" or "has trouble keeping up with the flow of a normal

conversation". That is, the poor fit could be due to mistargetting of items to children (i.e., a poor spread of items across the full range of their scores), or that the data could have floor (or ceiling) effects due to poor discrimination of items among younger children (Hagquist et al. 2009).

In the Rasch analysis, 35 items were excluded from the set of 65 because of local dependency, lack of fit, and disordered thresholds, all of which could contribute to the lack of adequate fit of the hypothesized unidimensional factor structure. Local dependency means that some items were highly correlated with other items and led to a lack of fit based on tests of residuals (person and item). There were disordered thresholds in 11 items, indicating that the response scale was not functioning as it should and that the meaning of the responses for those items was unclear.

Using Rasch analysis, we also showed that the 65-item SRS could be reduced to a 30-item subset with good internal consistency using data from 4-year-olds. Examination of DIF for the 3 data points using the 30-item subset showed no evidence of uniform or non-uniform DIF. The 30-item subset was shown to be unidimensional and a well-fitting measurement model for the 30-item set, explaining about 88 % of the variance of the 65-item SRS. Post hoc examination of the 30-item subset using the full sample showed that the data were a good fit to the Rasch model and the internal consistency based on the PSI was also high. Examination of concurrent validity of the 30-item subset total score with child outcome measures showed that the total score was positively associated with CBCL subscales and RBS-R domains, indicating that severity of autistic social impairment was associated with severity of internalizing/externalizing behaviour and repetitive behaviours. Similar to the findings of Constantino and Gruber (2005), there was a negative association of the SRS with adaptive functioning (VABS-II Adaptive Behavior Composite score) but no statistically significant relationship with language skills (PLS-4 Total score).

The 30-item subset may prove useful for research in preschool samples, as it is easier to implement and yields a single-dimension construct (as proposed by Constantino and Gruber 2005). We were also able to show the utility of Rasch analysis as a complementary approach in examining the measurement properties of the SRS and in the refinement of the SRS. Indeed, although the 65-item SRS did not perform well statistically, possibly because of the age of the children, the 30-item subset appeared to represent "markers" of autistic social impairment that functioned well across age groups, even in the narrow follow-up interval ranging from 12 to 36 months. The 30-item subset also meets the assumptions underlying the Rasch model and therefore may have potential for use in evaluating the severity of autistic social impairment as a single dimension in other clinical samples of preschool children. These results suggest the need for the findings be replicated in larger independent samples with wider age ranges. Independent studies should also evaluate the psychometric properties and clinical utility of the 30-item subset.

One limitation of this study is that the SRS was designed for use in 4- to 18-year-olds, yet our data came from preschoolers at the lower end of the age range (4 years). Another limitation was that fewer than 100 4-year-olds were available for analyses at any single time point. We were able to test 205 4-year-olds in this study by creating an accelerated longitudinal dataset by combining data from three time points. As a consequence, age (or cohort effects) could influence our results.

In conclusion, our findings suggest that the structure of the 65-item SRS cannot be described as unidimensional in this sample of 4-year-olds with ASD. Moreover, a substantial number of SRS items functioned poorly in not discriminating well among preschool children with ASD, at least in this sample. The complementary Rasch analysis showed that the 65-item SRS could be reduced to a 30-item subset with little loss in explanatory power, with the relationships between the 30-item subset and other outcome measures being similar to those found with the 65-item SRS. Use of this subset of SRS items is therefore recommended in measuring the severity of social impairment among preschool age children with ASD.

Acknowledgments This study was supported by the Canadian Institutes of Health Research, Autism Speaks, the Government of British Columbia, Alberta Innovates—Health Solutions, and the Sinneave Family Foundation. The authors thank all the families who participated in the *Pathways in ASD* study. The authors also acknowledge the members of the *Pathways in ASD Study Team*. These members had equal contribution to the study and are listed here alphabetically: Liliana Abruzzese, Megan Alexander, Susan Bauld, Ainsley Boudreau, Colin Andrew Campbell, Mike Chalupka, Lorna Colli, Melanie Couture, Bev DaSilva, Vikram Dua, Miriam Elfert, Lara El-Khatib, Lindsay Fleming, Kristin Fossum, Nancy Garon, Shareen Holly, Stephanie Jull, Karen Kalynchuk, Kathryne MacLeod, Preetinder Narang, Julianne Noseworthy, Irene O'Connor, Kaori Ohashi, Jennifer Endre Olson, Sarah Peacock, Teri Phillips, Sara Quirke, Katie Rinald, Jennifer Saracino, Cathryn Schroeder, Cody Shepherd, Rebecca Simon, Mandy Steiman, Richard Stock, Benjamin Taylor, Lee Tidmarsh, Larry Tuff, Kathryn Vaillancourt, Stephen Wellington, Isabelle Yun, and Li Hong Zhong.

Appendix

See Table 5.

Table 5 30-item subset of SRS and original subscales

Item	Subscale	Description
SRS1	SMot	Fidgety in social situations than when alone
SRS4	AMan	Under stress, shows rigid or inflexible patterns of behavior
SRS5	SCog	Doesn't recognize when others take advantage
SRS6	SMot	Would rather be alone
SRS9	SMot	Clings to adults
SRS14	AMan	Not well coordinated
SRS15	SCog	Understands meaning of people's tone and facial expressions
SRS16	SCom	Avoids eye contact
SRS19	SCom	Gets frustrated trying to get ideas across in conversations
SRS20	AMan	Shows unusual sensory interests
SRS22	SCom	Plays appropriately
SRS23	SMot	Does not join group activities
SRS24	AMan	Difficulty with changes in routine
SRS25	SAw	Doesn't seem to mind being out of step
SRS29	AMan	Regarded by other children as odd
SRS30	SCog	Becomes upset in a situation with lots of things going on
SRS31	AMan	Can't get mind off something
SRS33	SCom	Socially awkward
SRS36	SCom	Difficulty relating to adults
SRS42	SCog	Overly sensitive to sounds
SRS44	SCog	Doesn't understand how events relate to one another
SRS46	SCom	Serious facial expressions
SRS49	AMan	Does extremely well at a few tasks,
SRS52	SAw	Knows when talking too loud
SRS56	SAw	Walks between two people talking
SRS57	SCom	Teased a lot
SRS58	SCog	Concentrates on parts rather than seeing the whole picture
SRS63	AMan	Touches in unusual way
SRS64	SMot	Too tense in social settings
SRS65	SMot	Stares or off into space

AMan autistic mannerisms, *SMot* social motivation, *SCog* social cognition, *SCom* social communication, *SAw* social awareness

References

Achenbach, T. M., & Rescorla, L. A. (2000). *Manual for the ASEBA preschool forms & profiles*. Burlington, VT: University of Vermont, Research Center for Children, Youth, & Families.

American Psychiatric Association (2000). *Diagnostic and statistical manual of mental disorders DSM-IV-TR* (4th ed.; text revision). Washington, DC: American Psychiatric Association.

Andrich, D. (1988). *Rasch models for measurement: Sage University Paper Series on Quantitative Measurement in the Social Sciences*. Newberry Park, CA: Sage.

Andrich, D. (2004). Controversy and the Rasch model. *Medical Care, 42*(1 Suppl), i7–i16.

Andrich, D., Sheridan, B., & Luo, G. (2007). *RUMM2020: Rasch unidimensional measurement models software, version 4.1*. Perth, Western Australia: RUMM Laboratory.

Bodfish, J. W., Symons, F. J., Parker, D. E., & Lewis, M. H. (2000). Varieties in repetitive behavior in autism. *Journal of Autism and Developmental Disorders, 30*, 237–243.

Centers for Disease Control and Prevention. (2012). Prevalence of autism spectrum disorders—Autism and developmental disabilities monitoring network, 14 sites, 2008. *Morbidity & Mortality Weekly Report Surveillance Summaries, 61*(3), 1–19.

Charman, T. (2007). Autism and its impact on child development. In R. E.Tremblay, R. G. Barr, & R. D. V. Peters (Eds.). *Encyclopedia on early childhood development* [online] (pp. 1–5). Montreal, Quebec: Centre of Excellence for Early Childhood Development. Retrieved November 8 2011 from http://www.child-encyclopedia.com/documents/CharmanANGxp_rev.pdf.

Chawarska, K., Paul, R., Klin, A., Hannigen, S., Dichtel, L. E., & Volkmar, F. (2007). Parental recognition of developmental problems in toddlers with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 37*, 62–72.

Constantino, J. N., & Gruber, C. P. (2005). *Social responsiveness scale*. Los Angeles, CA: Western Psychological Services.

Constantino, J. N., Gruber, C. P., Davis, S., Hays, S., Passante, N., & Przybeck, T. (2004). The factor structure of autistic traits. *Journal of Child Psychology and Psychiatry, 45*, 719–726.

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*(1), 2–11.

Frazier, T. W., Youngstrom, E. A., Kab, C. S., Sinclair, L., & Rezai, A. (2008). Exploratory and confirmatory factor analysis of the Autism Diagnostic Interview-Revised. *Journal of Autism and Developmental Disorders, 38*, 474–480.

Georgiades, S., Szatmari, P., Zwaigenbaum, L., Duku, E., Bryson, S., Roberts, W., et al. (2007). Structure of the autism symptom phenotype: A proposed multidimensional model. *Journal of the American Academy of Child and Adolescent Psychiatry, 46*, 188–196.

Hagquist, C., & Andrich, D. (2004). Is the sense of coherence-instrument applicable on adolescents? A latent trait analyses using Rasch-modelling. *Personality and Individual Differences, 36*, 955–968.

Hagquist, C., Bruce, M., & Gustavsson, J. P. (2009). Using the Rasch model in nursing research: An introduction and illustrative example. *International Journal of Nursing Studies, 46*, 380–393.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*, 1–55.

Lord, C., Risi, S., Lambrecht, L., Cook, E. H., Leventhal, B., DiLavore, P. C., et al. (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*, 205–233.

Lord, C., Rutter, M., & Le Couteur, A. (1994). Autism Diagnostic Interview-Revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *Journal of Autism and Developmental Disorders, 24*, 569–685.

Mirenda, P., Smith, I., Vaillancourt, T., Duku, E., Georgiades, S., Szatmari, P., et al. (2010). Validating the Repetitive Behavior Scale-Revised in young children with autism spectrum disorder. *Journal of Autism and Developmental Disorders, 40*, 1521–1530.

Muthén, L. K., & Muthén, B. (2008a). *Mplus 5.1 for Windows*. Los Angeles, CA: Author.

Muthén, L. K., & Muthén, B. (2008b). *Mplus user's guide*. Los Angeles, CA: Author.

Newschaffer, C. J., Croen, L. A., Daniels, J., Giarelli, E., Grether, J. K., Levy, S. E., et al. (2007). The epidemiology of autism spectrum disorders. *Annual Review of Public Health, 28*, 235–258.

Risi, S., Lord, C., Gotham, K., Corsello, C., Chrysler, C., Szatmari, P., et al. (2006). Combining information from multiple sources in the diagnosis of autism spectrum disorders. *Journal of the American Academy of Child and Adolescent Psychiatry, 45*(9), 1094–1103.

Sparrow, S. S., Cicchetti, D. V., & Balla, D. A. (2005). *Vineland II: A revision of the Vineland adaptive behavior scales: I. survey/ caregiver form*. Circle Pines: American Guidance Service.

Tennant, A., & Conaghan, P. G. (2007). The Rasch measurement model in rheumatology: what is it and why use it? When should it be applied, and what should one look for in a Rasch paper? *Arthritis and Rheumatism, 57*(8), 1358–1362.

Volden, J., Smith, I. M., Szatmari, P., Bryson, S., Fombonne, E., Mirenda, P., et al. (2011). Using the preschool language scale, fourth edition to characterize language in preschoolers with ASD. *American Journal of Speech-Language Pathology, 20*, 200–208.

Wright, B. D. (1977). Solving measurement problems with the Rasch model. *Journal of Educational Measurement, 14*(2), 97–116.

Zimmerman, I., Streiner, R., & Pond, R. (2002). *Preschool language scale* (4th ed.). San Antonio, TX: Psychological Corp.