Brief Report: Concurrent Validity of Autism Symptom Severity Measures

**Stephanie S. Reszka · Brian A. Boyd ·
Matthew McBee · Kara A. Hume · Samuel L. Odom**

Abstract The autism spectrum disorder (ASD) diagnostic classifications, according to the DSM-5, include a severity rating. Several screening and/or diagnostic measures, such as the autism diagnostic and observation schedule (ADOS), Childhood Autism Rating Scale (CARS) and social responsiveness scale (SRS) (teacher and parent versions), include an assessment of symptom severity. The purpose of this study was to examine whether symptom severity and/or diagnostic status of preschool-aged children with ASD (N = 201) were similarly categorized on these measures. For half of the sample, children were similarly classified across the four measures, and scores on most measures were correlated, with the exception of the ADOS and SRS-P. While the ADOS, CARS, and SRS are reliable and valid measures, there is some disagreement between measures with regard to child classification and the categorization of autism symptom severity.

S. S. Reszka (✉) · B. A. Boyd
Department of Allied Health, Division of Occupational Science and Occupational Therapy, University of North Carolina, 321 S. Columbia Street, Bondurant Hall CB #7122, Chapel Hill, NC 27599-7122, USA
e-mail: stephanie.reszka@unc.edu

Present Address:
M. McBee
East Tennessee State University, Johnson City, TN, USA

M. McBee · K. A. Hume · S. L. Odom
Frank Porter Graham Child Development Institute, University of North Carolina, Chapel Hill, NC, USA

Introduction

The proposed changes to the forthcoming diagnostic and statistical manual of mental disorders, DSM-5 (http://www.dsm5.org) would include severity criteria for the autism spectrum disorders (ASD) category. This new criteria would combine autism disorder, Asperger syndrome, and pervasive developmental disorder—not otherwise specified (PDD-NOS) into one larger ASD category. As a result of this collapse, reliable and valid measurement of autism severity will be even more important in the determination of services for children with a diagnosis of ASD (Matson et al. 2012).

Currently, the *Childhood Autism Rating Scale* (CARS; Schopler et al. 1986) and Social Responsiveness Scale (SRS; Constantino 2002) are two commonly used measures that include a symptom severity estimate. Previously, higher raw scores on the autism diagnostic and observation schedule (ADOS; Lord et al. 1999) indicated the presence of more deficits that are characteristic of individuals with ASD, suggesting a greater level of impairment, but the raw scores were not normalized to indicate severity (Gotham et al. 2009). A recent calibrated severity metric provides estimations of ASD symptom severity using ADOS scores (see Gotham et al. 2009). Generally, severity is measured in several areas for children with ASD: language delay, cognitive functioning, and behavioral issues (Gotham et al. 2009), however these are not necessarily considered the core features of ASD. Each of these measures, the CARS, SRS, and ADOS utilizes slightly different methods of evaluating the severity of ASD symptoms and have varied diagnostic cut-offs along the ASD spectrum.

The primary purpose of this study was to examine whether children's symptom severity and/or diagnostic status were similarly categorized across the four measures.

The two study goals were to examine: (1) the concurrent validity of the ADOS, CARS, and SRS (parent and teacher versions) and (2) the categorization of children's diagnostic status and symptom severity.

Methods

Data for this study were collected on 201 children as part of a larger study comparing the efficacy of school-based, comprehensive treatment models for preschoolers with ASD. Data were collected across four states (CO, NC, FL, and MN), and at the beginning of the school year. For each child, all measures were collected within a 6-week time window.

Participants

Children

At enrollment, the mean child age was 3.59 years ($SD = 0.56$, range 2.24–5.04). Most participating children were male (83.3 %) and ethnically non-Hispanic (64.6 %). In terms of racial status, 5.1 % were identified as Asian, 12.1 % were Black, 78.3 % were White, and 4.0 % were multiracial. To be eligible for the larger study, each child was required to have a clinical or school diagnosis of autism, PDD-NOS, or Asperger's Syndrome, or meet the autism spectrum cut-off score on the ADOS and Social Communication Questionnaire (SCQ; Rutter et al. 2003). If the child had an educational label of developmental delay (DD) instead of ASD, which is consistent with federal and state policy for children in this age range, then s/he must have met diagnostic criteria on both the ADOS and SCQ to be eligible for the study. It was not the point of our study to diagnose children, but rather screen them for potential eligibility and a DD educational label is reflective of the real-world heterogeneity when recruiting children through local school systems. The other study measures included the following: (1) Mullen Scales of Early Learning (Mullen 1995), which is a measure of children's cognitive and motor development. Trained research staff administered the visual reception, fine motor, expressive language, and receptive language subscales to the child. The mean standard score on the Mullen was 64.40 (N = 193, $SD = 19.6$, range 49–136). And (2) Preschool Language Scale, fourth edition (PLS-4; Zimmerman et al. 2003), which is a measure of children's auditory comprehension and expressive communication skills. The mean standard score on the PLS-4 was 68.23 (N = 198, $SD = 68.23$, range 50–134).

Parents

Most participating parents were female (88.2 %), non-Hispanic (66.8 %). Additionally, 5.2 % were identified as Asian, 13.0 % were black, 78.7 % were white, and 3.1 % were multiracial. Household annual income ranged from less than $20,000 (12.8 %) to over $100,000 (26.7 %). Parents completed the parent version of the SRS (SRS-P).

Teachers

Teachers completed the teacher version of the SRS (SRS-T). Participating teachers were almost exclusively female (98.6 %) and non-Hispanic (83.6 %), and identified themselves as white (97.3 %), with the remaining 2.7 % identifying themselves as black. most held a master's degree (56.2 %), while 37 % had a bachelor's, 2.7 % had an associate's, and 4.1 % had a degree above the master's level.

Diagnostic and Severity Measures

The measures examined in this study included the ADOS, CARS, and SRS parent and teacher versions. Both the ADOS and CARS were administered by trained and reliable project staff. The ADOS was administered by a research-trained and/or research reliable staff member at each site, and staff across sites met reliability criterion on a series of CARS training tapes prior to administration.

The ADOS is a semi-structured assessment of children's communication, social, and play skills. Module 1 is for children who are non-verbal or who have a few words. Module 2 is for children with phrase speech, while Module 3 is intended for children who are verbally fluent. In accordance with the suggested severity ratings, ADOS severity scores of 4–5 indicated autism spectrum disorder and scores from 6 to 10 indicated autism (Gotham et al. 2009). In this sample, 125 children were administered Module 1 of the ADOS, 57 were administered Module 2, and 15 were administered Module 3, while 4 children had missing data for the ADOS.

Using the CARS, the child is rated on 15 subscales based on observation (during the Mullen administration, in this case). To ensure consistency in CARS scoring across study sites and classrooms, the measure was completed based on observations of children's behavior during the structured administration of the Mullen and 15 min of unstructured time post-Mullen administration. The CARS includes items on socialization, communication, emotional response, and sensory issues. Each of the 15 items is rated on a scale from 0 to 4, with 4 indicating severe impairments. A CARS cutoff raw score of 25.5 was used to indicate autism spectrum disorder, with raw scores over 30 indicating autism (Chlebowski et al. 2010). The original CARS was used, as opposed to the newly released CARS2 (Schopler et al. 2010), because the CARS2 only became publicly available after the study was already underway.

This study used the original CARS, which is aligned with the currently available CARS2-ST, for children younger than 6 years of age.

The SRS is a 65-item rating scale that was completed by parents and teachers. The SRS provides information about children's social functioning including social awareness, social information processing, social reciprocal communication, social anxiety/avoidance behaviors, and stereotypic behavior/restricted interests. Each item is rated on a scale of 1 (not true) to 4 (almost always true). T-scores (mean of 50, standard deviation of 10) were used in the analyses, with a T-score of 60–75 indicating mild to moderate symptoms of ASD, and scores over 75 indicating severe symptoms. The SRS was normed with T-scores for parent and teacher versions, with separate norms within each for child gender. The appropriate scoring norms were used for each measure, as specified by the SRS manual. The preschool version of the SRS was used for children aged 36–47 months, and the standard version was used for children 48 months and older.

Results

Autism diagnostic and observation schedule scores ranged from 2 to 10, with a mean of 7.19 ($SD = 1.64$) suggesting that children in the sample tended to score in the milder end of the ASD category, but represented the full range of severity across the spectrum. The mean score on the CARS was 33.37 ($SD = 7.31$) with a range of 15–55.5. Similarly to the ADOS mean score, the mean score of 33.37 corresponds to the autism category for the CARS. The SRS-Teacher (SRS-T) version and SRS-Parent (SRS-P) versions both showed mean scores in the mild to moderate symptom category (66.27 and 73.70, respectively). Descriptive information for each measure is available in Table 1.

Question 1: Concurrent Validity at Pretest

The ADOS severity scores were significantly correlated with the CARS total score ($r = 0.432$, $p < 0.001$) and the total score on the teacher version of the SRS ($r = 0.418$, $p < 0.001$). The ADOS severity scores were not significantly correlated with scores on the SRS-P ($r = 0.088$, $p = 0.236$). The CARS was significantly correlated with

both versions of the SRS ($r = 0.558$, $p < 0.001$ for the teacher version; $r = 0.292$, $p < 0.001$ for the parent version). The SRS-Teacher and SRS-Parent scores were significantly correlated ($r = 0.275$, $p < 0.001$). The correlation matrix for these measures is shown in Table 2.

Question 2: Categorization of Diagnostic Status/ Severity

Nearly 98 % of the children scored on the spectrum according to the ADOS. The CARS scores classified 64.7 % of children as being on the spectrum. The SRS-Teacher and SRS-Parent scores classified 76.6 and 82.1 % of children as being on the spectrum, respectively. Diagnostic classification charts for each measure are available in Fig. 1.

A summary of children's diagnostic classifications across all measures is available in Table 3. Ratings were collapsed so that a score of 0 indicated that the child did not score on the autism spectrum, while a score of 1 indicated that a child would score in the autism spectrum range (mild/moderate/severe autism symptoms). As shown, for 92 cases (50 % of the sample) children were classified similarly across all measures. For another 25 cases (13.59 % of the sample), children were classified similarly on the ADOS and both versions of the SRS, but not the CARS. The remaining children scored on the spectrum on one or more of the measures. Almost 14 % scored on the spectrum according to the ADOS and both SRS versions, but not the CARS, followed by 10.33 % on the ADOS and SRS-Parent only. Another 6.52 % of children scored on the spectrum on the ADOS, CARS, and SRS-Parent. Approximately 6 % scored on the spectrum on both the ADOS and SRS-Teacher and another nearly 6 % on the ADOS, CARS, and SRS-Teacher. Almost 4 % scored on the spectrum only on the ADOS. Just over 2 % scored on the spectrum only on the ADOS and CARS. Finally, 1 % of children scored on the spectrum according to the SRS-Parent and SRS-Teacher forms only, and 0.54 % scored on the spectrum only on the SRS-Parent. For approximately 76 % of the sample (140 cases), children were similarly classified on at least three of the four measures.

Table 1 Descriptives for measures

Measure	N	Mean (SD)	Range
ADOS severity	198	7.19 (1.64)	2.00–10.00
CARS total score	200	33.37 (7.31)	15.00–55.50
SRS-Teacher total score	200	66.27 (9.66)	42.00–90.00
SRS-Parent total score	185	73.70 (14.27)	42.00–111.00

Table 2 Bivariate correlations of measures

	ADOS severity Score	CARS total score	SRS-Teacher
CARS total score	0.432 (<.001)	–	–
SRS-Teacher total score	0.418 (<.001)	0.558 (<.001)	–
SRS-Parent total score	0.088 (.236)	0.292 (<.001)	0.275 (.001)

p values in parentheses

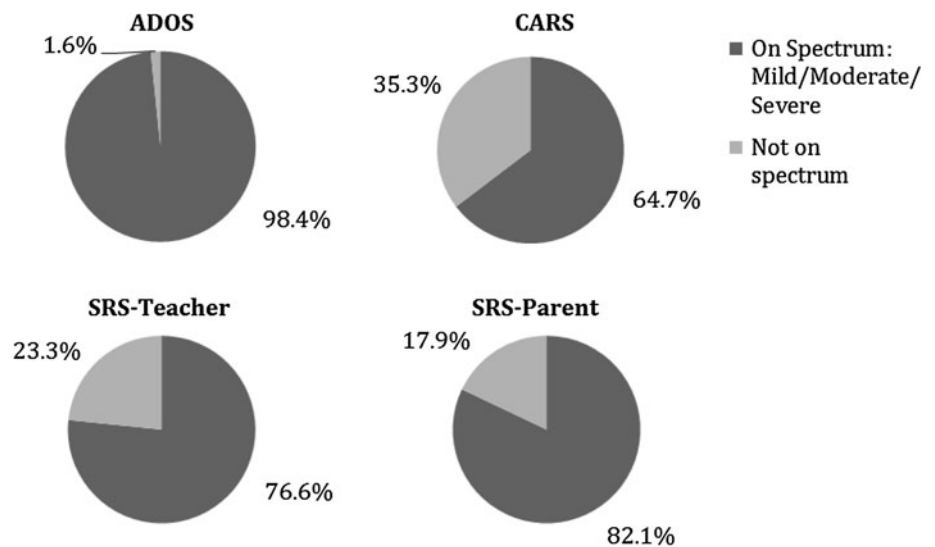**Fig. 1** Diagnostic classification pie charts by measure

Table 3 Collapsed summary of diagnostic ratings

ADOS	CARS	SRS-Teacher	SRS-Parent	N	%
0	0	0	1	1	0.54
0	0	1	1	2	1.09
1	0	0	0	7	3.80
1	0	0	1	19	10.33
1	0	1	0	11	5.98
1	0	1	1	25	13.59
1	1	0	0	4	2.17
1	1	0	1	12	6.52
1	1	1	0	11	5.98
1	1	1	1	92	50.00
				184	100.00

17 cases were missing and not included in the analysis. 0 = not autistic/not on spectrum, 1 = on spectrum/mild autism/severe autism

Discussion

Generally, children's severity scores on the measures were correlated, indicating that the severity of autism symptoms was rated similarly across all measures, with the exception of the ADOS and SRS-Parent version. There were moderate to strong correlations between the CARS and all other measures, and between the SRS-T and all other measures. The ADOS was moderately correlated with both the CARS and SRS-T, but not with the SRS-P. Research suggests that scores on the SRS agree with clinical diagnosis a significant portion of the time and the SRS teacher and parent versions have shown correlations ranging from 0.75 to 0.91 in a clinical sample (Constantino et al. 2003), while this sample showed a weaker, but still significant, correlation of 0.275. Interestingly, the parent version of the SRS was correlated, albeit moderately, with all other measures with the exception of the ADOS. However, the statistical significance of some of the more modest correlations may be an artifact of the relatively large sample size used in this study.

The differences in ADOS and SRS-Parent scores seen in this study may reflect potential variations in child behaviors across different contexts; all measures except the SRS-P were completed in the school context, while the SRS-P reflects parental views of child behaviors at home. It is important to consider the context under which these measures of symptom severity were collected. The parent measures were not always correlated with measures taken in the school context by teachers or research staff, and children may display different behaviors at home than they would in a classroom or research setting. Thus the context may be a factor in potential disagreements between parents' and clinicians' or practitioners' interpretations of symptom severity or autism diagnosis.

For half of the sample, children were similarly classified across all measures. About three quarters (76 %) of the sample were similarly classified on at least three of the four measures. Ratings on the CARS appear to be the most conservative regarding diagnosis, as only 64.7 % (119 children) were rated as having an ASD diagnosis using the CARS, while nearly all (98.4 %; 181 out of 184) of the children were rated as having a diagnosis on the spectrum according to the ADOS. However the ADOS, along with the SCQ, was used to determine children's study eligibility, and was selected because it is considered a gold-standard measure for ASD diagnosis.

While the children in this study were between the ages of 3 and 5, previous research comparing the ADOS and CARS for diagnosing toddlers with ASD suggests that there is a significant agreement between the two for diagnosing ASD in toddlers, matching clinical judgment (Ventola et al. 2006). Children in this study tended to have

mild to moderate symptoms of autism. The CARS is better at diagnosing children who tend to be lower functioning than those who are higher functioning (Mayes et al. 2009), which may explain some of the discrepancy between CARS classification and the other measures. A newly released version of the CARS (CARS2-HF) assesses verbally fluent, more high-functioning children, but currently is only available for children age 6 and older.

The proposed changes to the DSM-5 include severity criteria for the ASD category, allowing ratings of symptoms along "a continuum from mild to severe rather than a simple yes or no diagnosis to a specific disorder" (APA 2012). Given these changes, measures of symptom severity may become more critical in autism research and clinical practice. While the severity measures used in this study may not match the severity criteria in the proposed DSM-5, this study is a first step toward examining the agreement, or lack thereof, of commonly used measures of autism symptom severity. Additional future studies should examine the relationships between the current measures of severity described in this study with the severity classifications that will be found in the DSM-5.

While there are instruments that can produce reliable and valid assessments of autism severity available, this study demonstrates that there is some disagreement among several of these measures with regard to child classifications and the categorization of symptom severity. The type of measure used could affect child classifications, and by extension, services provided to these children.

References

American Psychiatric Association (2012). *DSM-5 proposed criteria for autism spectrum disorder designed to provide more accurate diagnosis and treatment.* http://www.dsm5.org/Documents/12-03%20Autism%20Spectrum%20Disorders%20-%20DSM5.pdf. Accessed 12 Mar 2013.

Chlebowski, C., Green, J., Barton, M., & Fein, D. (2010). Using the cars to diagnose ASD. *Journal of Addiction, 40*, 787–799.

Constantino, J. N. (2002). *Social Responsiveness Scale (SRS)*. Los Angeles: Western Psychological Services.

Constantino, J. N., Davis, S. A., Todd, R. D., Schindler, M. K., Gross, M. M., Brophy, S. L., et al. (2003). Validation of a brief 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.

Gotham, K., Pickles, A., & Lord, C. (2009). Standardizing the ADOS scores for a measure of severity in autism spectrum disorders. *Journal of Autism and Developmental Disorders, 39*, 695–705.

Lord, C., Rutter, M., DiLavore, P., & Risi, S. (1999). *Autism diagnostic observation schedule (ADOS)*. Los Angeles, CA: Western Psychological Services.

Matson, J. L., Beighley, J., & Turygin, N. (2012). Autism diagnosis and screening: Factors to consider in differential diagnosis. *Research in Autism Spectrum Disorders, 6*, 19–24.

Mayes, S. D., Calhoun, S. L., Murray, M. J., Morrow, J. D., Yurich, K. K. L., Mahr, F., et al. (2009). Comparison of scores on the checklist for autism spectrum disorder, childhood autism rating scale, and Gilliam Asperger's disorder scale for children with low functioning autism, high functioning autism, Asperger's disorder, ADHD, and typical development. *Journal of Autism and Developmental Disorders, 39*, 1682–1693.

Mullen, E. (1995). *The Mullen Scales of early learning*. Circle Pines, MN: American Guidance Service.

Rutter, M., Bailey, A., & Lord, (2003). *Social communication Questionnaire (SCQ)*. Los Angeles: Western Psychological Services.

Schopler, E., Reichler, R. J., & Renner, B. R. (1986). *The Childhood Autism Rating Scale (CARS)*. Los Angeles, CA: Western Psychological Services.

Schopler, E., Van Bourgondien, M. E., Wellman, G. J., & Love, S. R. (2010). *Childhood autism rating scale, second edition (CARS2)*. Los Angeles: Western Psychological Services.

Ventola, P. E., Kleinman, J., Pandey, J., Barton, M., Allen, S., Green, J., et al. (2006). Agreement among four diagnostic instruments for autism spectrum disorder in toddlers. *Journal of Autism and Developmental Disorders, 36*, 839–847.

Zimmerman, I., Steiner, V., & Pond, R. (2003). *Preschool Language Scale-IV*. San Antonio: Psychological Corporation.