2018

Use of a video scoring anchor for rapid serial assessment of social communication in toddlers

Natasha Marrus
Washington University School of Medicine in St. Louis

Stefanie Kennon-McGill
Washington University School of Medicine in St. Louis

Brooke Harris
Veteran's Administration Palo Alto Health Care System

Yi Zhang
Washington University School of Medicine in St. Louis

Anne L. Glowinski
Washington University School of Medicine in St. Louis

See next page for additional authors

Follow this and additional works at: https://digitalcommons.wustl.edu/open_access_pubs

Authors

Natasha Marrus, Stefanie Kennon-McGill, Brooke Harris, Yi Zhang, Anne L. Glowinski, and John N. Constantino

Video Article

Use of a Video Scoring Anchor for Rapid Serial Assessment of Social Communication in Toddlers

Natasha Marrus[1], Stefanie Kennon-McGill[1], Brooke Harris[2], Yi Zhang[1], Anne L. Glowinski[1], John N. Constantino[1]

[1]Department of Psychiatry, Washington University School of Medicine

[2]Veteran's Administration Palo Alto Health Care System

Correspondence to: John N. Constantino at constantino@wustl.edu

Abstract

Reciprocal social behavior (RSB), an early-emerging capacity to engage in social contingency—which is foundational for both social learning and social competency—is hypothesized to be disrupted in autism spectrum disorder (ASD). The ability to quantify the full range of RSB during the toddler period, when core symptoms of ASD often arise, is pivotal for evaluating early risk for ASD, characterizing social development, and tracking response to early interventions. However, important parameters of variation in RSB—especially prior to the development of verbal language—can be nuanced and difficult to characterize using questionnaire-based methods. To address this challenge, we developed a system for measuring quantitative variation in RSB in toddlers (ages 18 - 30 months) that incorporated not only standard questionnaire data from caregivers but also a novel set of video-referenced items, through which a respondent compares the behavior of a subject to that observed in a short video of a young child manifesting a highly competent level of social communication. Testing of this measure in a general population sample of twins confirmed that both the video-referenced items and the RSB Total Score (video-referenced items plus non-video-referenced items) displayed unimodal, continuous distributions, strong internal consistency, marked preservation of individual differences, and extremely high heritability. In addition, video-referenced items were particularly sensitive to quantifying incremental changes in social communication, a major element of RSB, over the course of early childhood development. Scores on the vrRSB clearly differentiated children with and without ASD and these data comprise an initial validation of this promising method for quantifying early RSB—cross-sectionally, over time, and as a function of early intervention.

Video Link

The video component of this article can be found at https://www.jove.com/video/57041/

Introduction

A prerequisite for the development of social competency is reciprocal social behavior (RSB), the normative capacity to engage in emotionally appropriate, socially contingent behavior with others[1]. Disruption of RSB, which is associated with marked deficits in social communication, interpersonal awareness, and social interest and engagement, is a defining feature of autism spectrum disorder (ASD). This heritable neurodevelopmental condition, which presents with a broad range of impairment in social and adaptive function, generally arises by the toddler period and can respond to early intervention. Quantification of RSB in early life is thus relevant both for clinical assessment of ASD risk, as well as characterization of the typical development of social competency.

The Social Responsiveness Scale (SRS) is a rapid quantitative rating scale of RSB applicable for children over age 30 months. This caregiver-report measure capitalizes on respondents' cumulative observations of a child in naturalistic situations to obtain nuanced ratings of RSB. Items query raters' impressions about the extent of a child's social communication deficits and restricted repetitive behaviors, the two core symptom domains of ASD[2], thereby indexing quantitative deficits in RSB. Extensive research with the SRS has confirmed its strong psychometric properties[3], including a continuous, unimodal score distribution which encompasses variation in the general population while also differentiating individuals with ASD[4,5]. Quantitative variation in RSB exhibits longitudinal stability[6], consistent with a behavioral trait; is distinguishable from adaptive function[7], another salient aspect of impairment in ASD; and is heritable[5,8,9,10,11], supporting the relevance of RSB to the genetics and neurobiology underlying ASD.

Because the SRS only extends downward to 30 months of age, an alternative rapid, quantitative measure of RSB for toddlers is needed to a) quantify early signs of ASD risk and b) track responses to early interventions for ASD. However, measurement of RSB in toddlers is complicated by substantive inter-individual variation in social communicative abilities, which can be difficult for raters to differentiate reliably based on written descriptions. Furthermore, caregivers' impressions may be susceptible to reduced consistency, due to the inherent subjectivity of interpreting limited verbal output in toddlers, the sensitivity of toddler behavior to situational and environmental contexts, and the accelerated progression of social abilities during a prolific period of development.

To address these methodological challenges, we developed the video-referenced Rating of Reciprocal Social Behavior (vrRSB), a quantitative index of RSB designed for children 18 - 30 months of age[12]. We used two strategies to represent the trait-based continuity of RSB while conferring sensitivity to early developmental change in RSB. First, to promote homology between the vrRSB and the SRS, we adapted a subset of items from the SRS so that they were developmentally appropriate for toddlers. Second, we incorporated a series of novel, unique video-referenced items in which caregivers rated their child's social communicative behavior in comparison to that of a videotaped child serving as a scoring anchor. This approach capitalized on the frame of reference effect, in which providing additional contextual information to raters improves both reliability and validity of behavioral trait measures[13,14]. Here, a visual frame of reference was provided by a 3-minute video montage of a typically developing 19-month old displaying strong social communication skills while engaging in a play-based interaction with an adult. Since aspects of RSB are highly correlated within individuals[15,16], the video featured a multifaceted view of a single child, rather than brief clips of several children, to enhance ecological validity. Previous studies additionally support the ecological validity of the video-referenced approach, as subtle signs of ASD can be detected on videotaped observations prior to clinical diagnosis[17,18]. Selected video segments visually highlighted hallmarks of social communication, including turn-taking, motivation to engage, and responsiveness to social and emotional cues[1], thereby ensuring content validity for essential aspects of RSB, and spoken language use was limited to avoid rater misattribution of language ability to RSB.

To our knowledge, there is currently no alternative brief measure of RSB for toddlers which indexes dimensional variation in RSB in both the general population and cases of ASD or other social delays. Established caregiver-report screeners for ASD and related delays in toddlers, such as the Modified Checklist for Autism in Toddlers (M-CHAT)[19] or the Communication and Symbolic Behavior Scales (CSBS)[20] generally consist of items indicating whether or not a behavioral milestone has been achieved, and are not explicitly designed to quantify subtle variation in social competency. Additionally, unlike the vrRSB, these instruments are not specific to RSB and do not index levels of core autistic symptoms (*i.e.,* impairment in social communication and restricted, repetitive behavior). Observational assessments, such as the Autism Diagnostic Observation Schedule[21], have been developed to assist with diagnosis; however, these are generally time-intensive, with a duration of 40 - 60 minutes, and require trained raters, limiting their feasibility in some clinical settings. No training is required to complete the vrRSB, making it appropriate for use with caregivers, service providers, and other informants who have sufficient recent experience observing a child's naturalistic social interactions and who understand written English. The measure, including the 3-minute video, is brief, requiring 15 - 20 minutes, and is amenable to administration in-person or remotely, as it can be mailed to participants or completed online[22]. Thus, the vrRSB's brevity, accessibility, and flexibility facilitate its implementation in research, clinical, and educational settings to obtain serial, quantitative measurement of RSB at an earlier age than previously possible.

Protocol

All methods described here have been approved by the Institutional Review Board of Washington University in St. Louis (Human Resource Protection Office Number: 201208010).

1. Obtaining the vrRSB

1. To obtain a copy of the vrRSB, send an email request to the senior author, John Constantino (constantino@wustl.edu).
2. Specify a preference for the online version of the vrRSB or the hard-copy format.
    1. Please note that the online version ("Online_vrRSB") is formatted for implementation in REDCap, a widely used, secure, online data collection tool (www.projectredcap.org) available to individuals at institutions which are members of the REDCap Consortium.
    2. Please note that the hard-copy version includes PDF's of the questionnaire, a copy of the password-protected online link to the video clip, and scoring instructions. Interested parties may also request paper copies of these files, as well as a DVD with the video.

2. Uploading and Enabling the Online vrRSB in REDCap

1. Save the online file for the vrRSB (Online_vrRSB) in a known computer location.
2. Before uploading the **Online_vrRSB** into REDCap, create a new project to house the questionnaire in REDCap. First, log in to REDCap and click **New Project** in the ribbon at the top of the page. This step will launch a **Create a New REDCap Project** page.
3. Enter study-specific text into the fields **Project title** and **Purpose of the project**. These fields are required to proceed.
4. For the prompt **Start project from scratch or begin with a template?** select the radio button **Upload a REDCap project xml file**. A window providing access to computer files should then appear; choose the **Online_vrRSB** file from step 2.1 and click **Open**.
5. Scroll to the bottom of the page and click **Create Project**. This step will open a **Project setup page**, which provides a workflow for incorporating surveys into a REDCap project.
6. In the box labeled **Main Project Settings**, click **Enable** next to the option **Use surveys in this project**. This step allows the vrRSB to be accessible within the project.
7. Proceed to the next box, labelled **Design your data collection instruments & enable your surveys**. Select **Online designer**. This will open a screen with a table of instruments within the project.
8. In the row which contains the vrRSB, click on **Enable a survey**. A **Set up my survey** tab should then appear.
9. Scroll to the bottom of the page and click **Save changes** to complete the enabling of the vrRSB in a REDCap project.

3. Customizing the Presentation of the Online vrRSB in REDCap

1. Click **Project Home** in the sidebar menu and locate the **Quick Tasks** box.
2. In the **Quick Tasks** box, select **Online designer and upload data dictionary** to open a Data collection instruments table listing the vrRSB as part of the REDCap project.

3. To modify instructions within the vrRSB text, click the **vrRSB** listing. Click on the pencil in the upper left of specific text boxes to edit existing text and hit **Save** after each modification.
    1. In the box labeled **Welcome page**, provide study-specific background information as appropriate (for suggestions, see step 4.2).
    2. Click **Add fields** to allow entry of additional participant characteristics if needed. The current template includes age of the rated child in months, sex, and the relationship of the rater to the child.
    3. Edit the instructions preceding the video-referenced items by inserting a study contact person in the event of technical difficulties with the video.

4. To customize survey format (*e.g.* background color, font), click **Survey settings** in the row for the vrRSB within the data collection instruments table.
5. To obtain an online link for the survey, return to **Project Home** and click on **Manage survey participants** in the **Quick Tasks** box. A **Public survey link** page will open, which provides a URL for the survey. Click the button next to the URL to copy the survey link.
    1. To email a link to study participants, select the **Participant list** tab. Click **Add participants** to insert emails and **compose survey invitations** button to generate an introductory email.

4. Administering the vrRSB

1. Select respondents who understand written English and who have spent at least 6 hours observing the target child (18 - 30 months of age) in the past month. Successful completion of the vrRSB does not require or assume specialized training, only that respondents have routinely observed the behavior of a young child in his or her natural interactions with other people. Respondents may include caregivers, clinicians, other professionals who have worked closely with a child, or researchers who wish to use the video scoring anchor to generate calibrated ratings to compare multiple children.
    NOTE: A 30-minute observation of a child engaged in a semi-structured interactive play session has been shown to be sufficient for raters to provide scores distinguishing children with and without ASD[12]. The 6-hour benchmark is intended to account for greater expected variability in real-world social situations.
2. When presenting the survey to respondents, provide background information for clarification, *e.g.*, the purpose for administering the vrRSB, the use of video-referenced items, and the expected time for completion (10 - 15 minutes).
    1. For cases in which respondents participate remotely, include such information in a cover letter accompanying the vrRSB.
    2. For cases in which participants complete the vrRSB online, provide the web address to access the vrRSB.
3. Explain whether the video of the scoring anchor is to be accessed online or from a DVD. For online access, include the web address and password.
4. Inform respondents whom to contact for help should they have difficulty accessing the full video, seeing the images, or hearing sound. In the online version, this notification is included in the instructions preceding the video-referenced items (step 3.3).
5. Remind respondents to attend to the entire 3-minute video clip, which they may watch more than once.
    NOTE: Respondents next complete questions 1 - 13, the video-referenced items, in which they are to compare the subject child's current social abilities to that of the child in the video - the scoring anchor. Following completion of the video-referenced items, respondents are to complete the remaining non-video-referenced items, which describe characteristic aspects of social behavior. Both video-and non-video-referenced items use a four-point Likert rating scale, and written instructions precede both sets of items in the body of the questionnaire (**Appendix 1**).

5. Automated Scoring of the Online vrRSB

1. To view an individual's scores (which may be useful in clinical settings), click **Add/edit records** under **Data Collection** in the sidebar menu.
    1. In the following screen, under **Choose an Existing Participation Identification #**, select the participant from the dropdown menu. The Record Home Page for that participant will appear, which contains a table listing instruments collected for that participant.
    2. To access total and subscale scores for that participant, click **Status** in the row for **vrRSB scores**.

2. To view vrRSB scores for all participants in a given study, return to the project home page and click **Create report**.
    1. Enter a name for the report in the text box and proceed to step 2, which allows the selection of data fields to include in the report.
    2. Next to the option **Add all fields from selected instruments**, select **vrRSB scores** from the dropdown menu to obtain total and subscale scores. For item-level data, select **vrRSB**. To keep only a subset of available fields in the report, click the red **x** to the right of unwanted fields so they are excluded.
    3. Scroll to the bottom of the page and click **Save report**. Accept the option to **Return to my reports and exports**, which will display a table including a row with the newly created report.
    4. Click **View report** to see scores (**Table 1**) for all participants. To return to the table of viewing options, click the 'My Reports and Exports' tab.
    5. Click **Export data** and select the format of choice (*e.g.,* csv format). Then click the export data button at the lower right. In the follow-up screen, click on the icon confirming the data format to continue.
    6. Click the **"Stats and charts"** button to view descriptive statistics (*e.g.,* means for total scores, standard deviations) and scatterplots of vrRSB scores in REDCap.

6. Manual Scoring of the vrRSB

1. Identify items as forward-scored or reverse-scored (**Table 1**, Classification of Forward- and Reverse-Scored Items).

NOTE: Because the scale is designed to quantify deficits in RSB, an indicator of the level of core autistic symptoms, higher point values indicate lower levels of RSB. Forward-scored items are those for which a more affirmative answer indicates less appropriate reciprocal social behavior. Reverse-scored items are those for which a more affirmative answer indicates more appropriate reciprocal social behavior.

2. Assign each response a value, as indicated in **Table 1** (Point Values for Item Ratings).
3. Assign a value of zero for missing item responses. Surveys with more than four missing responses are considered invalid.
4. Once items have been assigned the proper point values, add the appropriate items together for the main score(s) or subscale score(s) of choice (**Table 1**).
   NOTE: Main scoring options include the RSB Total Score (video-referenced plus non-video-referenced RSB items) and the Video-referenced Score; subscale item totals are also noted for the 2 core symptom domains for ASD (*i.e.,* social communication and restricted, repetitive behavior). Scores ≥2 standard deviations below the mean, in the bottom 5% of the scores obtained in a general population sample[12], are suggestive of deficiencies in RSB (**Table 1**).
5. Note that some items are not included in scoring for the vrRSB. These supplemental items provide general information on aspects of development other than RSB, which may be of interest to developmental researchers. See the caption of **Table 1** for further details.

Representative Results

The vrRSB was tested in 156 twin pairs epidemiologically ascertained from the Missouri Family Register as part of the Early Reciprocal Social Behavior Study, a prospective, longitudinal study of the development of RSB[12]. We selected a twin study design to examine whether the vrRSB could detect individual differences in RSB within toddlers from the general population and whether this variation was heritable. Both these features have been demonstrated with the SRS in older children[5,8] and are key for evaluating the vrRSB's ability to capture genetically informative variance in RSB which is relevant for both typical and atypical development.

Caregivers completed behavioral measures on the twins at 18, 24, 36, and 48 months of age; the vrRSB (primarily the hard-copy version, as the online version was developed later in the study) was collected at the 18- and 24-month time points. The sample was primarily Caucasian (80.1%) and non-Hispanic (91.6%), similar to the Missouri population[23], whereas median annual household income of the sample ($60,001 - $90,000) was higher than that of Missouri ($47,200). Informal feedback from participants regarding the survey was generally positive, with over 95% of returned surveys being correctly completed. To account for non-independence of twin data, analyses of reliability, behavioral correlations, and longitudinal differences included data from one twin selected at random. One child, who underwent a developmental regression and ultimately received a community diagnosis of ASD, was excluded from these twin analyses.

Both the Video-referenced Scores and the RSB Total Score demonstrated a continuous, unimodal distribution at 18 and 24 months across the general population sample (**Figure 1**). For the Video-referenced Score, a greater preponderance of lower scores at 24 months of age suggested a slight floor effect due to more advanced development (higher scores indicate less RSB). Distributions of the RSB Total Score, like those for the SRS3, were right-skewed, reflecting the small proportion of individuals with significant deficits in RSB. Similar score distributions have been observed in an independent sample of primarily Hispanic, bilingual twins from California (unpublished data, Harris, B., 2017).

Good internal reliability for both video-referenced items and the RSB Total score was observed at 18 months (Cronbach's alpha Video-referenced Score = .841, p <.001; Cronbach's alpha RSB Total Score = .879, p <.001) and 24 months (Cronbach's alpha Video-referenced Score = .881, p <.001; Cronbach's alpha RSB Total Score = .868, p <.001). The strong internal consistency suggests that these item indices represent cohesive constructs throughout toddlerhood.

Test-retest reliability was calculated according to intraclass correlation coefficients (ICCs; two-way mixed single measures) based on scores obtained at the 18- and 24-month time points. Good 6-month test-retest reliability was observed for the Video-referenced Score (ICC(3,1) = .505, p <.001, **Figure 2a**), and strong test-retest reliability was observed for RSB Total Score (ICC(3,1) = .735, p <.001; **Figure 2b**).

In addition to demonstrating test-retest reliability and, by extension, the capacity to detect trait-related consistency in RSB across time, an important goal for the vrRSB was to track progress in RSB over development. Both the Video-referenced and RSB Total Score were able to detect significant improvement in RSB over time, manifested by decreased scores (lower scores signify better RSB) between 18 and 24 months (**Table 2**). Video-referenced Scores showed a larger effect size for these decreases vs. the RSB Total Score (Cohen's d = 0.66 vs. 0.48). Separate calculations for the non-video-referenced items (**Table 2**), which were not designed to track change, revealed a much smaller effect size for these items (Cohen's d = 0.22).

Since variation in RSB has been shown to be heritable in older populations, twin-twin concordances were calculated for Video-referenced Scores and RSB Total Scores at 18 months (**Table 3**). Higher intraclass correlations in monozygotic (identical) twins, who share all their genetic material, vs. dizygotic (fraternal) twins, who share on average half of their genetic material, are consistent with heritability of RSB. Concordances for monozygotic twins were quite high and in many cases roughly double those of dizygotic twins, supporting a strong role for genetic influences. Of note, monozygotic twin concordances also indicate of the upper limit of possible measurement error. For both vrRSB scores, the possible measurement error is less than 9% at the level of the entire twin sample (error = 1 - monozygotic twin concordance).

As a test of concurrent validity, we examined correlations of scores on the M-CHAT, a screener for ASD, with the vrRSB. Positive correlations were consistent with the vrRSB's capacity to capture variation related to ASD risk (Video-referenced Score: r = .444, p <.001; RSB Total Score: r = .654, p <.001). We also obtained correlations between vrRSB scores and expressive vocabulary (percentile scores) as measured by the MacArthur-Bates Communicative Development Inventory Short Form[24], given an anticipated relationship between language and social communication. Significant correlations were found for both the Video-referenced Score (r = -.419, p <.001) and the RSB Total Score (r = -.309, p <.001).

Lastly, in pilot testing (**Table 4**), we compared scores for twins without a diagnosis of ASD and a small sample of children with either a community diagnosis of ASD or a rating of autism or autism spectrum on the Autism Diagnostic Observation Schedule[21] (mean age(SD): 24.75(5.56) months; age range: 17 - 33 months). Both Video-referenced Scores and RSB Total Scores differentiated between children with and without ASD, with an effect size (Cohen's d) of -1.39 for the video-referenced score and -1.66 for the RSB Total Score. Group differences were also observed on item scores for ASD core symptom domains (Social Communication: Cohen's d = -1.58; Restricted, Repetitive Behavior: Cohen's d = -1.52). At the level of quantitative RSB scores, both the 18-month Video-referenced Score and RSB Total Score were correlated with total scores on the 36-month SRS (Video-referenced Score: r = .298, p = .001; RSB Total Score: r = .624, p <.001).
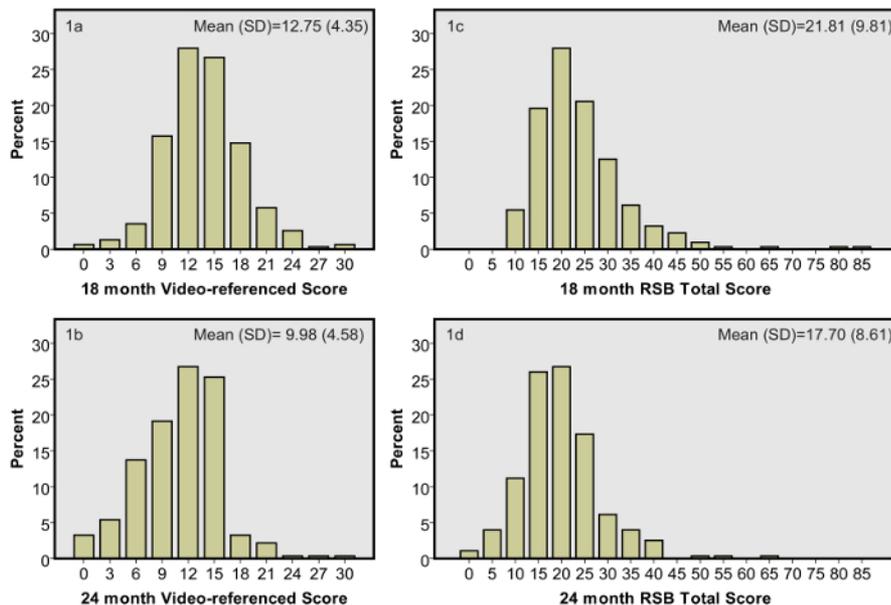
Figure 1: Continuous Unimodal Score Distributions for Video-referenced Score and RSB Total Score at 18 and 24 months
Panels illustrate scores for the vrRSB, including both Video-referenced Score and RSB Total Score, which is comprised of video-referenced items plus traditionally presented survey items adapted from the Social Responsiveness Scale. These distributions are based on a sample of twins ascertained from the general population (n = 156 pairs). Panels 1a and 1b show scores for video-referenced items at 18 and 24 months while panels 1c & 1d show RSB Total Scores at 18 and 24 months. In all cases, scores are continuously distributed and unimodal, consistent with RSB operating as a dimensional construct in toddlers. Means and standard deviations (SD) are indicated for each time point in upper right of panel. Please click here to view a larger version of this figure.
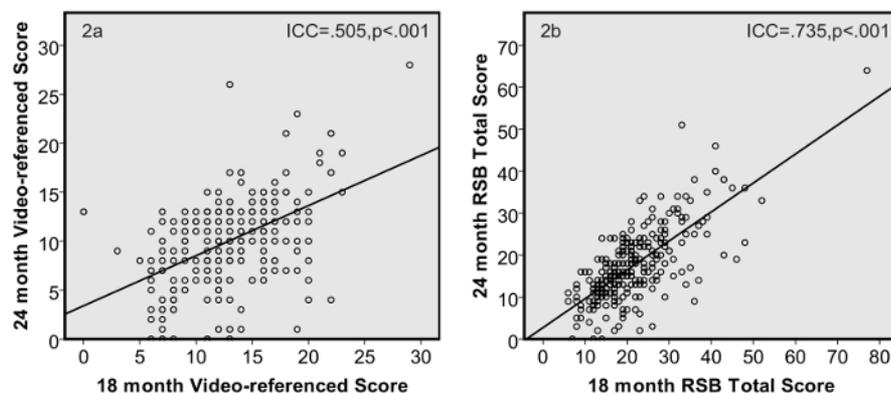
Figure 2: 6-month Test-Retest Reliability for Video-referenced Score and RSB Total Score
Scatterplots illustrate a cross-time comparison at 18 and 24 months for subjects' Video-referenced Score (2a) and RSB Total Score (2b). Intraclass correlation coefficients (ICC(3,1)) show good test-retest reliability for the Video-referenced Score and strong test-retest reliability for RSB Total Score. Please click here to view a larger version of this figure.

Classification of Forward- and Reverse-Scored Items	
Forward-Scored Items	15 - 16, 18 - 26, 28 - 34, 36, 40 - 41, 47, appendix item 1
Reverse-Scored Items	1 - 13; 14, 17, 27, 35, 37 - 39, 42

Point Values for Item Ratings							
	Video-referenced Items				*Non-video Referenced Items*		
Rating option	Not at all	Somewhat but less than child in video	About the same as child in video	More than child in video	Not true	Sometimes true	Often true	Almost always true
Forward-scored	None				0	1	2	3
Reverse-scored	3	2	1	0	3	2	1	0

Item Lists for Main Scores and Subscale Scores			
Scale/Subscale	*Item numbers*	*Score ≥2 SD of Mean*	
		18 months	24 months
RSB Total	1 - 13;14 - 42; 47; appendix item 1	41	35
Video-referenced	1 - 13	21	19
Social Communication	1 - 15; 17; 19; 21; 25 - 28; 30; 32 - 33, 35 - 40, 42, appendix item 1	36	32
Restricted Repetitive Behavior	16, 18, 20, 22 - 24, 29, 31, 34, 41, 47	7	5

Table 1: Guidelines for Video-referenced Rating of Reciprocal Social Behavior (vrRSB) Scoring
This table provides a breakdown of item types, point values, and item lists used for scoring. The vrRSB contains 2 main scores, the RSB Total Score and the Video-Referenced Score, and 2 subscale scores, Social Communication and Restricted Repetitive Behavior, which include items related to the 2 core symptom domains of ASD. These scores quantify the level of deficits in RSB and thus also indicate the burden of core autistic symptoms. Note that select items address aspects of general development, and these items are not included in scoring. Items 49 and 50 provide an indication of the achievement of two key language milestones for toddlers: the use of spoken language and combining words into phrases or sentences. Item 43 regards unusual sleep patterns, with greater endorsement suggesting a less typical developmental profile. For items regarding spatial abilities (44, 48) and intellectual curiosity (45, 46), greater endorsement corresponds to more advanced development.

Paired Samples	Paired Differences					T	df	Sig.
	Mean	Standard Deviation	Standard Error Mean	95% Confidence				
				Interval				
				Lower	Upper			
18m Video-referenced Score –	2.91	4.38	0.37	2.17	3.64	7.79	137	<.001
24m Video-referenced Score								
18m RSB Total Score –	4.30	6.48	0.55	3.21	5.40	7.80	137	<.001
24m RSB Total Score								
18m non-video items –	1.40	4.47	0.38	0.65	2.15	3.68	137	<.001
24m non-video-items								

Table 2: Differences in Video-referenced Rating of Reciprocal Social Behavior (vrRSB) scores from 18 to 24 months
Paired t-tests are used to calculate differences between scores at 18 and 24 months (n = 138). Score differences are for Video-referenced Scores, RSB Total Scores, and scores for the non-video referenced items (RSB Total Score - Video-referenced Score). vrRSB indicates Video-referenced Rating of Reciprocal Social Behavior. Df indicates degrees of freedom. Sig. indicates significance.

Twin Type (pairs)	Video-referenced	RSB Total
MZ (n = 54)	.928 (p <.001)	.912 (p <.001)
DZ (n = 95)	.388 (p <.001)	.218 (p <.001)
Male MZ (n = 28)	.894 (p <.001)	.914 (p <.001)
Male DZ (n = 30)	.348 (p = .028)	.286 (p <.059)
Female MZ (n = 26)	.956 (p <.001)	.878 (p <.001)
Female DZ (n = 29)	.660 (p <.001)	.369 (p <.023)
Opposite Sex (n = 36)	.301 (p = .035)	.062 (p = .358)

Table 3: Twin-Twin Concordances for Toddler Reciprocal Social Behavior at 18 months
Twin-twin concordances [ICC(3,1)] for Video-referenced Scores and RSB Total Scores at 18 months are shown for monozygotic (MZ) and dizygotic (DZ) twins. Values are reported for all MZ and DZ twin pairs grouped together as well as twin pairs analyzed separately by gender. Seven unclassified pairs were not able to be included in these analyses.

	no-ASD (n=310)	ASD (n=14)	Mean Difference	T	Df	Sig.
	Mean (SD)	Mean (SD)				
Video-referenced	12.77 (4.44)	21.86 (8.09)	-9.08	-4.17	13.36	0.001
RSB Total	21.84 (9.79)	56.79 (28.17)	-34.95	-4.63	13.14	<.001
Social Communication	20.15 (8.14)	44.57 (20.16)	-24.42	-4.52	13.19	0.001
Restricted, Repetitive Behavior	1.68 (2.76)	12.21 (9.37)	-10.53	-4.20	13.10	0.001

Table 4: Comparison of Video-referenced Rating of Reciprocal Social Behavior (vrRSB) Scores for Toddlers with and without autism spectrum disorder (ASD)
Means, mean differences and results of t-tests for vrRSB scores are shown for comparison of scores between twins without an ASD diagnosis and a small group of children with ASD. SD indicates standard deviation. Df indicates degrees of freedom. Sig. indicates p-values for 2-tailed significant differences.

Appendix 1: Copy of the Video-Referenced Rating Scale of Reciprocal Social Behavior Please click here to download this file.

Discussion

The above example analyses demonstrate the vrRSB's strong psychometric properties as a quantitative toddler measure of RSB. Scores on the vrRSB showed several analogous features to the SRS, a well-validated measure of RSB for older children[3], including continuous score distributions in the general population, evidence of heritability, and the ability to distinguish children with and without ASD (in pilot testing for the vrRSB). vrRSB scores at 18 and 24 months of age correlated with 36-month SRS scores, confirming that trait-like stability of RSB is measurable in early childhood. Thus, the vrRSB represents an effective downward extension of the SRS and the first instrument, to our knowledge, to provide a dimensional measure of RSB quantifying individual differences in typically and atypically developing toddlers. This ability is in contrast to screeners like the M-CHAT and CSBS, which adopt a categorical approach to risk of ASD or other delays[19,20] and do not specifically index RSB. Implementation of the vrRSB may enhance the efficient delivery of early interventions, by facilitating rapid assessment of change in core ASD symptoms, which may show a distinct response compared to other areas of impairment in ASD, such as adaptive function. This tool also has the advantages of being brief, accessible, and amenable to multiple administration options, including the fully developed online option in REDCap, which are appropriate for both clinical and research settings. In addition to the Missouri study population reported here, our team has successfully administered this survey in a primarily Hispanic population of twins, with results showing comparable psychometric properties (unpublished data, Harris, B., 2017). This tool is thus applicable to diverse populations and would be easily amenable to translation into other languages, particularly as interpreting the scoring anchor's social behavior does not rely on spoken language.

We applied two approaches to overcome the inherent challenges of caregiver reporting for RSB in toddlers: a) adapting a subset of SRS items which could be made developmentally appropriate for toddlers and b) implementing a videotaped scoring anchor to provide a frame of reference for caregiver ratings. These strategies permitted a Video-referenced Score describing social communication, a central component of RSB, and a broader RSB Total Score covering social communication and restricted repetitive behavior, the two core symptom domains of ASD. Both scores demonstrated useful measurement properties. The Video-referenced Score, by using a video of a single reference child at all ages of assessment, was designed to be sensitive to change in RSB during toddlerhood, and of all scores, it showed the largest effect size for improvement in RSB from 18 to 24 months. This score was also more strongly correlated with expressive language than RSB Total Score, consistent with this scale's focus on social communication. In the case of the RSB Total Score, the incorporation of SRS items representing traits related to RSB and ASD was associated with stronger test-retest reliability and higher correlations with 36-month total SRS scores; these observations are in keeping with the enhanced sensitivity of Video-referenced Scores to incremental change in social communication. Both the Video-referenced and RSB Total Scores showed high monozygotic twin-twin concordances, implying low measurement error and strong reliability, while corroborating findings of good test-retest reliability, which may have been susceptible to underestimation given rapid developmental change between 18 and 24 months of age. Additionally, in pilot testing, Video-referenced and RSB Total Scores distinguished children with and without ASD. The effect size of this difference, while large for both scores, was slightly greater for RSB Total Score, which maximizes the opportunity to capture variation in autistic traits through a greater number of items and coverage of both ASD symptom domains

(i.e., restricted repetitive behavior in addition to social communication). Nevertheless, a strong effect size was also observed for the Video-referenced Score, suggesting its potential viability as a very brief, quantitative ASD screening measure in toddlers.

A critical step for the success of this protocol involves the appropriate selection of informants, who must have sufficient experience with the child to be capable of rating the child's level of RSB. To date, we have been successful with parents as informants, although this survey is also appropriate for other caregivers, such as daycare providers or other family members who have frequent contact with the child (*i.e.,* no less than 6 hours either interacting with the child or watching the child interact with others over the past month). Survey respondents should be reminded to watch the full video, so that the scoring anchor provides a uniform standard of comparison across raters, and it is helpful to inform respondents in advance to allow 15 - 20 minutes to view the video and complete the questionnaire. If the survey is administered online, recording the amount of time spent on the survey is helpful for quality control, since surveys completed too quickly raise concerns for validity. Finally, because the multiple steps for scoring the survey provide an opportunity for error, we recommend using an automated scoring algorithm and cross-checking sums of subscales with the total score.

One limitation of the vrRSB is the potential for floor effects with the video-referenced items at older ages, given that toddlers approaching 30 months are likely to display more advanced social communication than the 19-month old scoring anchor. Thus, at older ages, the RSB Total Score is expected to capture more representative variation in RSB. Additionally, because the scoring anchor displays strong social communication ability, some caregivers might be inclined to rate their child as similar to the scoring anchor across all items, rather than to consider differences between the child and the anchor. Our results, including the broad score distributions and sensitivity of video-referenced scores to developmental change, suggest this was not a major issue in this sample; nevertheless, subsequent iterations of the vrRSB may explore the use of multiple scoring anchors with varying levels of RSB, as well as reverse-scored video-referenced items to check for conceptual consistency of responses.

Future directions include testing of the vrRSB in larger samples of children with and without ASD, to confirm the measure's psychometric properties and to determine the extent to which vrRSB scores predict risk of ASD. A larger sample will also facilitate data reduction approaches to allow refinement of the instrument by removing any items with low contributions to variance in RSB. From the methodological standpoint, the promising psychometrics of the Video-referenced Score support the general utility of a video scoring anchor in promoting reliability and validity of a behavioral ratings. An infant scoring anchor could feasibly allow further downward extension of an RSB measure to younger toddlers and infants, which could enhance timeliness of referral to early interventions and broaden characterization of typical and atypical social developmental trajectories. In summary, the vrRSB provides important opportunities to improve clinical assessment of ASD risk and response to intervention. The vrRSB's ability to detect dimensional variation in RSB, change in RSB, and genetic influences on RSB at the stage when ASD emerges is also important for clarifying the early impact of genes and environment on RSB, which can guide future treatments for ASD and related disorders.

Disclosures

Acknowledgements

References

1. Constantino, J.N., Przybeck, T., Friesen, D., Todd, R.D. Reciprocal social behavior in children with and without pervasive developmental disorders. *J Dev Behav Pediatr.* **21** (1), 2-11 (2000).
2. American Psychiatric Association. *Diagnostic and statistical manual of mental disorders, fifth edition.* American Psychiatric Association. Arlington, VA. (2013).
3. Constantino, J.N. *Social Responsiveness Scale, 2nd Editions (SRS-2).* Western Pscychological Services. Torrance, CA. (2013).
4. Constantino, J.N., Hudziak, J.J., Todd, R.D. Deficits in reciprocal social behavior in male twins: evidence for a genetically independent domain of psychopathology. *J Am Acad Child Adolesc Psychiatr.* **42** (4), 458-67 (2003).
5. Constantino, J.N., Todd, R.D. Genetic structure of reciprocal social behavior. *Am J Psychiatry.* **157** (12), 2043-5 (2000).
6. Constantino, J.N. *et al.* Developmental course of autistic social impairment in males. *Dev Psychopathol.* **21** (1), 127-38 (2009).
7. Marrus, N., Underwood-Riordan, H., Randall, F., Zhang, Y., Constantino, J.N. Lack of effect of risperidone on core autistic symptoms: data from a longitudinal study. *J Child Adolesc Psychopharmacol.* **24** (9), 513-8 (2014).
8. Constantino, J.N., Todd, R.D. Autistic traits in the general population: a twin study. *Arch Gen Psychiatry.* **60** (5), 524-30 (2003).
9. Constantino, J.N., Zhang, Y., Frazier, T., Abbacchi, A.M., Law, P. Sibling recurrence and the genetic epidemiology of autism. *Am J Psychiatry.* **167** (11), 1349-56 (2010).
10. Lyall, K. *et al.* Parental social responsiveness and risk of autism spectrum disorder in offspring. *JAMA Psychiatry.* **71** (8), 936-42 (2014).
11. Virkud, Y.V., Todd, R.D., Abbacchi, A.M., Zhang, Y., Constantino, J.N. Familial aggregation of quantitative autistic traits in multiplex versus simplex autism. *Am J Med Genet B Neuropsychiatr Genet.* **150B** (3), 328-34 (2009).
12. Marrus, N. *et al.* Rapid video-referenced ratings of reciprocal social behavior in toddlers: a twin study. *J Child Psychol Psychiatry.* **56** (12), 1338-46 (2015).

March 2018 | 133 | e57041 | Page 8 of 9

13. Bing, M.N., Whanger, J.C., Davison, H.K., VanHook, J.B. Incremental validity of the frame-of-reference effect in personality scale scores: a replication and extension. *Journal of Applied Psychology.* **89** (1), 150-7 (2004).
14. Lievens, F., De Corte, W., Schollaert, E. A closer look at the frame-of-reference effect in personality scale scores and validity. *J Appl Psychol.* **93** (2), 268-79 (2008).
15. Constantino, J.N*. et al.* The factor structure of autistic traits. *J Child Psychol Psychiatry.* **45** (4), 719-26 (2004).
16. Frazier, T.W*. et al.* Confirmatory factor analytic structure and measurement invariance of quantitative autistic traits measured by the Social Responsiveness Scale-2. *Autism.* **18** (1), 31-44 (2014).
17. Ozonoff, S*. et al.* A Prospective Study of the Emergence of Early Behavioral Signs of Autism. *J Am Acad Child Adolesc Psychiatry.* **49** (3), 256-266.e1-2 (2010).
18. Werner, E., Dawson, G. Validation of the phenomenon of autistic regression using home videotapes. *Arch Gen Psychiatry.* **62** (8), 889-895 (2005).
19. Robins, D.L., Fein, D., Barton, M.L., Green, J.A. The Modified Checklist for Autism in Toddlers: an initial study investigating the early detection of autism and pervasive developmental disorders. *J Autism Dev Disord.* **31** (2), 131-44 (2001).
20. Wetherby, A.M., Allen, L., Cleary, J., Kublin, K., Goldstein, H. Validity and Reliability of the Communication and Symbolic Behavior Scales Developmental Profile With Very Young Children. *J Speech Lang Hear Res.* **45** (6), 1202-1218 (2002).
21. Lord, C., Rutter, M., DiLavore, PC, Risi S. *Autism Diagnostic Observation Scale.* Western Psychological Services. Los Angeles. (2000).
22. Wolff, J.J., Boyd, B.A., Elison, J.T. A quantitative measure of restricted and repetitive behaviors for early childhood. *J Neurodev Disord.* **8**, 27 (2016).
23. *Missouri QuickFacts from the US Census Bureau.* Retrieved from: <https://www.census.gov/quickfacts> (2013).
24. Fenson, L*. et al.* Short-form versions of the MacArthur Communicative Development Inventories. *Applied Psycholinguistics.* **21**, 95-116 (2000).