



Published in final edited form as:

J Child Psychol Psychiatry. 2015 December ; 56(12): 1338–1346. doi:10.1111/jcpp.12391.

Rapid video-referenced ratings of reciprocal social behavior in toddlers: A twin study

Natasha Marrus¹, Anne L. Glowinski¹, Theodore Jacob², Ami Klin^{3,4,5}, Warren Jones^{3,4,5}, Caroline E. Drain^{1,6}, Kieran E. Holzhauer¹, Vaishnavi Hariprasad¹, Rob T. Fitzgerald¹, Erika L. Mortenson¹, Sayli M. Sant¹, Lyndsey Cole¹, Satchel A. Siegel¹, Yi Zhang¹, Arpana Agrawal¹, Andrew Heath¹, and John N. Constantino¹

¹ Department of Psychiatry, Washington University, St. Louis, MO, USA

² Department of Psychology, Palo Alto University, Palo Alto, CA, USA

³ Marcus Autism Center, Children's Healthcare of Atlanta, Atlanta, GA, USA

⁴ Division of Autism & Related Disabilities, Department of Pediatrics, Emory University School of Medicine, Atlanta, GA, USA

⁵ Center for Translational Social Neuroscience, Emory University, Atlanta, GA, USA

⁶ Department of Neurology, Washington University, St. Louis, MO, USA

Abstract

Background—Reciprocal social behavior (RSB) is a developmental prerequisite for social competency, and deficits in RSB constitute a core feature of autism spectrum disorder (ASD). Although clinical screeners categorically ascertain risk of ASD in early childhood, rapid methods for quantitative measurement of RSB in toddlers are not yet established. Such measurements are critical for tracking developmental trajectories and incremental responses to intervention.

Methods—We developed and validated a 20-minute video-referenced rating scale, the video-referenced rating of reciprocal social behavior (vrRSB), for untrained caregivers to provide standardized ratings of quantitative variation in RSB. Parents of 252 toddler twins [Monozygotic (MZ)=31 pairs, Dizygotic (DZ)=95 pairs] ascertained through birth records, rated their twins' RSB at two time points, on average 6 months apart, and completed two developmental measures, the Modified Checklist for Autism in Toddlers (M-CHAT) and the MacArthur Communicative Development Inventory Short Form (MCDI-s).

Results—Scores on the vrRSB were fully continuously distributed, with excellent 6-month test-retest reliability ([intraclass correlation coefficient] ICC=0.704, $p<0.000$). MZ twins displayed markedly greater trait concordance than DZ twins, (MZ ICC=0.863, $p<0.000$, DZ ICC=0.231, $p<0.012$). VrRSB score distributions were highly distinct for children passing versus failing the M-CHAT ($t=-8.588$, $df=31$, $p<0.000$), incrementally improved from 18-24 months, and were inversely correlated with receptive and expressive vocabulary on the MCDI-s.

Correspondence, John N. Constantino, 660 S. Euclid Ave, Campus Box 8134, St. Louis, MO 63110, USA; constantino@wustl.edu.

Conflicts of interest statement: See Acknowledgements for disclosures.

The rest of the authors have declared that they have no potential or competing conflicts of interest in relation to this work.

Conclusions—Like quantitative autistic trait ratings in school-aged children and adults, toddler scores on the vrRSB are continuously distributed and appear highly heritable. These ratings exhibited minimal measurement error, high inter-individual stability, and developmental progression in RSB as children matured from 18-24 months, supporting their potential utility for serially quantifying the severity of early autistic syndromes over time and in response to intervention. In addition, these findings inform the genetic-environmental structure of RSB in early typical development.

Keywords

autism; reciprocal social behavior; video; twins; toddlers

Introduction

Reciprocal Social Behavior (RSB) refers to the capacity to engage in emotionally appropriate, socially contingent behavior with others. RSB encompasses social awareness, appropriate interpretation and response to interpersonal cues, and the motivation to interact with others (Constantino, Przybeck, Friesen, & Todd, 2000), all of which emerge in early childhood and serve as the underpinnings of effective social relatedness. Numerous studies have demonstrated that RSB is continuously distributed in the general population as early as pre-school age (Constantino, Hudziak, & Todd, 2003; Constantino, et al., 2000; Ronald et al., 2006; Skuse et al., 2009), and extreme disruptions in RSB comprise a primary deficit of the autism spectrum disorders (ASDs), clinical conditions associated with severe enduring impairments in adaptive functioning. Twin and family studies of individuals with and without ASDs have demonstrated that deficits in RSB are highly heritable (Constantino & Todd, 2000, 2003; Le Couteur et al., 1996; Piven, Palmer, Jacobi, Childress, & Arndt, 1997; Wolff, Narayan, & Moyes, 1988), highlighting the potential role of RSB as an endophenotype in the neurobiology underlying ASD and related neuropsychiatric disorders.

In this study, we explored a ‘downward extension’ in age of rapid quantitative characterization of RSB to cover the developmental period when a) the characterizing features of autistic syndromes are typically first appreciable in clinical settings; and b) specific early developmental interventions for autism are typically first implemented in public health practice. Although a number of well-established screening instruments for ASDs, including the M-CHAT (Pandey et al., 2008; Robins, Fein, Barton, & Green, 2001) and the Communication and Symbolic Behavior Scales Developmental Profile Infant-Toddler Checklist (Pierce et al., 2011; Wetherby, Brosnan-Maddox, Peace, & Newton, 2008) have proven capable of identifying clinical-level deficits or delays in RSB in this age range, ascertainment of change over time with screening instruments is often limited to specifying whether a milestone has been reached or a categorical deficit has been resolved. As has been the case for older children, quantitative ratings, standardized by developmental age and gender, might confer distinct advantages for tracking developmental trajectories and incremental responses to early intervention, as well as for the discovery of associations between the developmental course of RSB and other biobehavioral traits.

Research involving the Social Responsiveness Scale (SRS)--a brief 65-item caregiver questionnaire that operationalizes RSB (and associated deficits that characterize autistic syndromes)--has demonstrated the utility of a rapid, quantitative approach. The SRS capitalizes upon caregivers' cumulative observations of children in their natural social settings, so that RSB can be reliably quantified by untrained raters. Over the course of more than 100 studies involving over 20,000 children (Constantino, 2011; Frazier et al., 2014), implementation of the SRS has been shown to distinguish ASD-affected children from controls (difference on the order of 3 SD) index heritable components of the autistic syndrome that aggregate in unaffected family members of children with ASD (Constantino, Zhang, Frazier, Abbacchi, & Law, 2010; Lyall, 2014, in press; Virkud, Todd, Abbacchi, Zhang, & Constantino, 2009), and measure RSB in large populations, allowing for serial measurement of RSB in research, clinical, and educational settings.

Currently, the preschool version of the SRS can reliably capture inter-individual variation in RSB as young as 30 months (Constantino & Gruber, 2012); however, this lower limit is beyond the age interval when symptoms of ASD are most often first appreciable in clinical practice. Attempts to develop similar rating scales for younger ages have proven challenging, especially in preverbal children, for whom substantive inter-individual variations in social and communicative ability can be subtle and difficult for raters to differentiate reliably based on written descriptions.

Thus, to calibrate the measurement of RSB among children age 18-30 months, we implemented a 3-minute video montage as a scoring anchor or 'frame of reference' for caregivers to use when rating key features of their child's RSB. The 'frame of reference' effect, in which contextualization is included for items involving behavioral traits, has been demonstrated to enhance validity and reliability of behavioral ratings (Bing, Whanger, Davison, & VanHook, 2004; Lievens, De Corte, & Schollaert, 2008), and videotaped observations of young children have been used successfully to identify behaviors associated with ASDs (Ozonoff et al., 2010; Werner & Dawson, 2005). In this study, we used this videotaped scoring anchor, against which caregivers rated their children, to contextualize a series of questions about social communication, which were followed by non-video-referenced items related to characteristic features of autistic syndromes--and refer to the entire scale as the video-referenced rating of reciprocal social behavior (vrRSB).

Here, we explored the psychometric properties of the vrRSB in an epidemiological sample of toddler twins. We hypothesized that RSB measured with the vrRSB in toddlers would demonstrate 1) a continuous unimodal distribution similar to that for older children and adults in the general population, 2) patterns of intraclass twin correlations suggesting genetic influences, 3) intra-individual temporal stability, and 4) developmental progression between ages 18-24 months, a period of rapid social-communicative maturation. Support for these hypotheses would substantiate the potential for feasible implementation of quantitative ratings of RSB in toddlers to monitor response to developmental intervention and to quantify heritable elements of social competency that are characteristically impaired in the early stage of autistic syndromes.

Methods

Participants

252 twins and their families participated in the 1st wave of the Early Reciprocal Social Behavior Study (4D068479), a longitudinal study prospectively investigating the developmental course of RSB at ages 18, 24, 36, and 48 months. Twins were epidemiologically ascertained through the Missouri Family Register, a database of birth records providing a source population for two waves of data collection: 1) 359 twin pairs born between September-December 2011 and 2) 260 twin pairs born between August and December 2012. Families were recruited via a letter describing the study, followed by a phone call. Phone numbers were obtained using standard proprietary and commercial databases and by requesting in the original letter that families provide a phone number. For twins to be eligible, the consenting family member was required to be the legal guardian and primary caregiver of the twins and to speak fluent English (the sole spoken language in 93.9% of Missouri households ('Missouri QuickFacts from the US Census Bureau,' 2013)). Demographic and medical information was collected via phone interview.

Five hundred fifty seven recruitment letters were mailed to 619 families ascertained from the birth records. Eleven were ineligible due to the death of a twin (n=10) or adoption (n=1). Phone numbers could not be found for 132 families, and for an additional 140 families, there was either no answer or a depersonalized voicemail. In 88 cases, a message was left on a personalized voicemail or the family inquired about the study via mail or phone message, but after several attempts by the study team, insufficient contact was established for enrollment. In total, 204 families were contacted to discuss enrollment; 162 of these enrolled and 126 completed the first full wave of data collection. There were no appreciable differences between those enrolled and not enrolled on the basis of information in the birth records, with the exception of maternal education, as the enrolled sample had a higher proportion of parents with college and graduate degrees. Table 1 summarizes characteristics of the sample.

In addition to twin families, a small contrast sample of 6 children with ASD, aged 18-24 months, was also recruited through informational postings in doctors' offices and early childhood special education centers. Parents contacted the study coordinator to pursue enrollment. These children were diagnosed with ASD by a community physician, or were suspected of having ASD due to parental concern for social delays.

Informed consent was obtained from all participating families and all study procedures were approved by the WUSM Human Research Protection Office (IRB) and by the State of Missouri Department of Health and Senior Services Institutional Review Board.

Measures

Video-referenced Rating of Reciprocal Social Behavior (vrRSB)—The rating scale used to quantify RSB in this study (the vrRSB) contains an initial 'video-referenced' section, for which informants rated their a child's social-communicative behavior relative to a typically-developing child depicted in a 3 minute-video montage viewed immediately prior to completing questions. Parents had the option of viewing this video clip via DVD or a

password-protected link online. This videotaped child, who served as a ‘scoring anchor,’ was 19 months of age and was shown engaging in an interactive play session with an adult. Video segments were selected for displays of typical social developmental milestones highlighting hallmarks of RSB, including turn-taking, motivation to engage, and responsiveness to social and emotional cues (Figure 1). The goal of the scoring anchor was to reduce discrepant interpretations on items by providing informants with a common naturalistic standard for comparison. By design, the instrument implemented video footage of a single child as the scoring anchor. As extremely high intercorrelations for various traits of social competency have been demonstrated within individuals (Constantino et al., 2004; Frazier, et al., 2014), we reasoned that exhibiting multiple facets of a single child's behavior would enhance ecological validity by allowing raters to integrate a meaningful appraisal of capacity for reciprocal social behavior across a montage of behaviors displayed in the video sequence. The video is available via password-protected internet link (available from the corresponding author).

Thirteen video-referenced items solicit from the rater a comparison of the participant child's behavior to the child in the video; these comprise the first set of items in the instrument. An additional 31 non-video-referenced items reflecting quantitative aspects of social behavior (homologous to item constructs validated in prior research on autistic traits), and developmentally appropriate for ages 18-30 months, comprise the second section of the scale. All items were designed to tap the informants' recall of accumulated experience with the child in naturalistic settings and represented behaviors related to social communication (SC) or restricted, repetitive behaviors (RRB), core symptom domains of ASD per DSM 5 (American Psychiatric Association, 2013). Visual inspection of score distributions for each item showed that no items were dichotomously endorsed. Both the video-referenced items and the entire vrRSB exhibit strong internal consistency, with Cronbach's alpha of .877 and .931, respectively. To establish whether video-referenced ratings could index inter-individual differences in RSB relevant to ASD, the video-referenced items of the vrRSB were first validated among untrained, blinded raters (n=11) who scored archived videotapes of children with (n=18) and without confirmed research diagnosis of ASD (n=22), aged 18-24 months. Videotapes depicted preverbal children engaged in 30-minutes of structured, interactive play with an adult. Video-referenced ratings accurately distinguished children with and without ASDs ($t=-4.835$, $df=38$; $p<.000$) and exhibited very strong inter rater reliability.

Goldsmith Child Zygosity Questionnaire—This parent-report measure was administered over the phone to all families with same-sex twin pairs. It consists of 27 questions regarding the similarity of physical characteristics between twins, including the frequency with which one twin is mistaken for the other by parents, relatives, and strangers. Agreement between this questionnaire and DNA markers/blood type used to classify monozygosity vs. dizygosity is over 93% (Price et al., 2000).

Modified Checklist for Autism in Toddlers (M-CHAT)—The M-CHAT is a parent-report, 23-item screener of social developmental milestones with moderate sensitivity and

high positive predictive value for future ASD (Pandey, et al., 2008; Robins, et al., 2001). Missing 3 or more questions results in a failed score, which is predictive of ASD.

MacArthur-Bates Communicative Development Inventory Short Form (MCDI-s), Level 1—The MCDI-s (level 1 infant version) consists of a parent-report, 89-word checklist that measures receptive and expressive vocabulary. It has been normed in typically developing children ages 8-18 months and demonstrates strong concordance to standardized behavioral assessments and the extensively validated full length ‘Words and Gestures’ form (Fenson et al., 2000). The short form provides a more rapid format to facilitate data collection, while still assessing receptive and expressive language development. For children older than 18 months at baseline assessment, the percentile score recorded was in relation to the 18 month-old norms.

Analysis

Analyses were conducted in SPSS version 20 or SAS 9.3. Student t-tests and chi-square analyses were used to compare group differences. To account for the non-independence of twin data, we incorporated data from one twin per family selected at random, in linear regressions, correlations, and analyses of test-retest reliability and score differences based on gender or change over time. Outliers were excluded from these analyses. Intraclass correlation coefficients (ICCs) were used for twin-twin correlations and test-retest reliability and Pearson's correlations were used for bivariate relationships. In a linear regression analysis exploring sources of variance in vrRSB ratings, maternal education was an ordinal variable dummy coded for low (associate's degree or less), mid-level (4 year college degree), and high maternal education (master's degree or above), and percentile of expressive language on the MCDI-s was used for expressive vocabulary. A repeated measures ANOVA (PROC GLM, SAS 9.3) including these factors was conducted to examine the impact of increasing age on changes in scores from 18 to 24 months. Given the strong internal consistency for both video-referenced items and the entire vrRSB, where applicable, we present scores for the 13-item video-referenced subset as well as a ‘RSB Total Score’ generated from both video and non-video-referenced items.

Results

RSB ratings in toddlers are continuously distributed

Scores for both the video-referenced subset and the RSB total were fully continuously distributed, as shown for males and females in Figure 2 (video-referenced scores: kurtosis=.805, skew=.110; RSB total scores: kurtosis=5.550, skew=1.612). Higher scores indicate lower developmental capacity for RSB. Sex differences were observed for both video-referenced and RSB total scores (video-referenced scores: $t=2.815$, $df=124$, $p<.006$, RSB total scores: $t=3.348$, $df=124$, $p<.001$). Significant sex differences were particularly attributable to items describing social communicative behaviors (SC items: $t=3.343$, $df=124$, $p<.001$); trend-level differences were observed for items describing restricted, repetitive behaviors (RRB items: $t=1.652$, $df=124$, $p<.102$). A linear regression testing for effects of gender, ethnicity, maternal education, and expressive vocabulary on video-referenced and RSB total scores at 18 months (Table 2) found a significant effect of gender and expressive

vocabulary on both indices derived from vrRSB (video-referenced scores: $F(5,123)=5.72$, $r^2=.1951$, $p<.0001$, gender: $t=3.29$, $p<.0013$, expressive vocabulary: $t=-3.81$, $p<.0002$; RSB total: $F(5, 123)=4.69$, $r^2=.1657$, $p<.0006$; gender: $t=3.54$, $p<.0006$, expressive vocabulary: $t=-3.01$, $p<.0032$) No significant effects were observed for ethnicity.

vrRSB Scores Exhibit Strong Test-Retest Reliability and Improve over the Course of Early Development

ICCs were calculated for the subset of participants (one twin from $n=70$ pairs) whose parents completed the vrRSB at both 18 and 24 months. Six-month test-retest reliability for the entire instrument was similar to previous autistic trait measures, with an $ICC=0.704$, $p<.000$ when removing two outliers, one of whom was diagnosed with an ASD following parental report of social communicative *regression* between 18 and 24 months, as shown in Figure 3.

As a group, participants' scores improved significantly over the 6 month time period for video-referenced scores and RSB total scores (Table 3). Mean changes in all sub-scores, with the exception of the RRB scores, were in the expected direction of acquisition of improved social developmental competency over the age interval from 18 to 24 months. A repeated measures ANOVA exploring whether the decrease in scores between 18 and 24 months was accounted for by age—a proxy for developmental progression—demonstrated a significant within-subjects main effect of age for the video-referenced and RSB total scores (video-referenced: Wilks' lambda, $F(1,63)=932$, $p<0.0362$) and RSB Total: Wilks' lambda, $F(1,63)= 934$, $p<.0382$). Interaction terms for age and gender, ethnicity, maternal education were not statistically significant with $\alpha=0.05$.

Higher Scores on vrRSB Identify Children at Risk for ASD

Video-referenced and RSB total scores were compared for toddlers who passed ($n=222$) versus failed ($n=30$) the M-CHAT screen by total number of missed items. Mean scores for M-CHAT passes vs. failures were significantly different (video-referenced: $t=-3.304$, $df=250$, $p<.001$; RSB Total: $t=-8.588$, $df=31$, $p<.000$), with higher video-referenced and RSB total scores associated with failing screening items on the M-CHAT. Scores for SC items and restricted, RRB items were also significantly greater in children who failed the M-CHAT (SC: $t=-6.948$, $df=250$, $p<.000$); RRB: $t=-4.077$, $df=30$, $p<.000$). As higher vrRSB scores correspond to lower levels of RSB, these results preliminarily support the relationship between quantitative ratings using the vrRSB and categorical screening information derived from the M-CHAT. Additionally, video-referenced and RSB Total scores were examined for a contrast sample of toddlers with ASD or suspected ASD ($n=6$). As occurred for children who failed the M-CHAT, mean video-referenced scores and RSB total scores, as well as SC and RRB scores for these subjects, were significantly greater than for the general population twin sample (video-referenced: $t=-3.912$, $df=5$, $p<.011$; RSB Total: $t=-4.866$, $df=5$, $p<.005$; SC: $t=-4.579$, $df=5$, $p<.006$; RRB: $t=-4.991$, $df=5$, $p<.004$).

Twin Concordance for vrRSB Ratings Suggestive of Substantial Heritability

ICCs were calculated for all 126 twin pairs and compared between MZ and DZ twins (Table 4). MZ twins of both genders exhibited extremely high ICCs for both video-referenced

scores (Males: .863; Females: .960) and RSB total scores (Males: .879, Females: .779) (Table 4). These correlations were primarily attributable to SC items, which exhibited extremely high MZ twin-twin ICCs, in contrast to RRB items, which exhibited lower ICCs. Like-sex DZ twin pairs exhibited substantially lower correlations, and correlations for opposite-sex DZ pairs were distinctly less than like-sex DZ pairs (Like-sex, video-referenced: .497, RSB Total: .350; Opposite-sex, video-referenced: .358, RSB total: .093). A contrasting relationship between like-sex and opposite-sex DZ pairs has been previously reported for ASD traits (Constantino and Todd, 2003; Hallmayer et al., 2011), suggesting that gender may modulate these associations.

Level of RSB is Associated with Language Development

As deficits in social communication comprise a key feature of ASD, we asked whether early capacity for RSB demonstrated an association with receptive and expressive vocabulary on the MCDI-s. RSB total scores demonstrated a moderate inverse correlation for number of words understood, a proxy for receptive language development ($r = -.271$, $p < .002$), and number of words produced ($r = -.265$, $p < .003$), a proxy for expressive language development. Similar inverse correlations were also observed for video-referenced subscale scores when related to both receptive ($r = -.413$, $p < 0.000$) and expressive ($r = -.364$, $p < 0.000$) language measures.

Discussion

Here we describe the development and initial testing of a novel video-referenced measure of reciprocal social behavior (vrRSB) in a population of epidemiologically-ascertained twins between the ages of 18-24 months. Our results show that, as has been observed for quantitative autistic trait measurements among school-aged children and adults in the general population, vrRSB scores exhibit a continuous distribution, whether considering only video-referenced items or including items describing traits and symptoms characteristic of ASDs. The continuous distribution, together with high intra-individual temporal stability of RSB total scores, demonstrates that stable inter-individual differences in the capacity for RSB—a core developmental capacity impaired in ASD—appear readily measurable using this method in the 18-24 month age range, and exhibit a wide range of variation in typically-developing children. Scores were higher in males than females, similar to descriptions of RSB in older children (Constantino and Todd, 2003), further supporting continuity in the ability to quantify this construct at younger ages. Correlational analysis in our twin population preliminarily indicate that scores on the vrRSB suggest a level of heritability similar to what has been observed for RSB in general populations of school-aged children and for ASD itself. Continued data collection will encompass sample sizes large enough to implement structural equation modeling for specifically estimating heritability among general population twins.

Response rate for this novel measure among participating families was high, 76%, and comparable to the more traditional questionnaires used in this study, demonstrating the feasibility of an internet-based, video-referenced standard for caregivers to use in serial ratings of their children's developmental progress. In addition to implications for heritability,

the very high MZ twin correlation (contrasting with a substantially lower DZ twin correlation) and the six-month stability results, place encouraging upper boundaries on measurement error, despite the difficulties inherent in ascertaining nuanced variations in the social behavior of very young children. Of note, ICCs for the video-referenced items in most analyses were generally greater than the ICCs for RSB total scores (Table 3), consistent with the benefit of a ‘frame of reference effect’ (Bing, et al., 2004; Lievens, et al., 2008) in enhancing the precision of measurement.

A major potential application for a quantitative measure of RSB in toddlers is to track severity of autistic impairment and to monitor incremental response to developmental therapies and early intensive behavioral intervention. While scores on the vrRSB were highly correlated within individuals across two time points 6 months apart, the mean scores for social communication derived from this instrument significantly decreased over time, suggesting group-level developmental maturation in the capacity for RSB, and the ability of the instrument to ascertain incremental progress in the acquisition of RSB over the course of early development. These initial results are encouraging with respect to the possibility that norms for the capacity for RSB at the successive mental/developmental ages could be established in order to identify when subtle social deficits suggestive of ASD first emerge in an individual.

Quantitative variation in RSB also exhibited an impressive association with clinical-level affectation in pilot testing involving known ASD cases, in relation to failed scores on the M-CHAT, and in a small pilot sample of children with early diagnoses of ASD. Among the two outlying vrRSB scores, one participant, whose scores markedly increased from 18-24 months, was clinically determined to have undergone an autistic regression. As expected, variation in early RSB also correlated with concurrent variation in receptive and expressive vocabulary on the MCDI-s. However, the relatively modest correlation observed suggests that while the MCDI-s is commonly used in clinical and public health settings, more sophisticated, albeit less practical language measures, might yield higher correlations, underscoring the potential clinical utility of monitoring RSB to index response to early interventions for developmental delay. We note that these initial data also exhibited distinct sexual dimorphism in the pace of acquisition of developmental competency for reciprocal social behavior inferred from the measurements, which is particularly important given the pronounced sex ratio observed in the incidence of clinical autistic syndromes (Constantino and Charman, 2012). Future analyses involving larger sample sizes and performance-based language assessments throughout the period of verbal language acquisition will help clarify the relationships between RSB, sex, and language development.

Although a major strength of the study involves the analysis of early RSB in a genetically informative twin sample, a limitation is that our twin sample measurements, comprising a small number of MZ twins, are too small to meaningfully implement more rigorous statistical models to estimate heritability. The evidence for heritability of RSB in toddlers is therefore preliminary. Test-retest analyses in larger populations will help to further establish the reliability of video-referenced ratings. Study recruitment is ongoing, with provisions for enhancing the diversity and representativeness of the current sample, including a parallel data collection in a population enriched for bilingual Hispanic twins, and the addition of

contrast group with ASD. Acquiring data at later ages through the longitudinal study design will importantly allow more in-depth characterization of the developmental course of RSB and the specificity of the vrRSB in predicting social outcomes.

Conclusion

In summary, video-referenced ratings of reciprocal social behavior offer a promising assessment option for quantifying inherited components of RSB in toddlers that represent critical targets for early identification and intervention. Use of such a measure, by promoting prompt recognition of whether or not an early intervention is leading to desired improvement in social competency, could inform more responsive, individualized modifications than are typically possible in clinical, educational, and public health settings. In research settings, such measurements—which appear to relate to heritable, sexually dimorphic elements of variation in early social competency—would be expected to enhance statistical power to explore the associations between behavioral manifestations of early social development and underlying genetic and neural mechanisms.

Acknowledgments

(NIH/NICHHD P30 HD062171, JNC PI). We would like to thank all the families participating in the Early Reciprocal Social Development study for their generous investment of time and effort. J.C. receives royalties from Western Psychological Services for the commercial distribution of the Social Responsiveness Scale, a quantitative measure of autistic traits for ages 30 months through adulthood.

References

- American Psychiatric Association. Diagnostic and statistical manual of mental disorders: DSM-5. American Psychiatric Association; Washington, D.C: 2013.
- Bing MN, Whanger JC, Davison HK, VanHook JB. Incremental validity of the frame-of-reference effect in personality scale scores: a replication and extension. *Journal of Applied Psychology*. 2004; 89(1):150. [PubMed: 14769127]
- Constantino JN. The quantitative nature of autistic social impairment. *Pediatric Research*. 2011; 69(5 Pt 2):55R–62R. 3086844.
- Constantino JN, Charman T. Gender bias, female resilience, and the sex ratio in autism. *Journal of the American Academy of Child and Adolescent Psychiatry*. 2012; 51(8):756–758. [PubMed: 22840545]
- Constantino, JN.; Gruber, CP. *Social Responsiveness Scale. Second Edition (SRS-2)*. Western Psychological Services; Torrance, CA: 2012.
- Constantino JN, Gruber CP, Davis S, Hayes S, Passanante N, Przybeck T. The factor structure of autistic traits. *Journal of Child Psychology and Psychiatry and Allied Disciplines*. 2004; 45(4):719–726.
- Constantino JN, Hudziak JJ, Todd RD. Deficits in reciprocal social behavior in male twins: evidence for a genetically independent domain of psychopathology. *Journal of the American Academy of Child and Adolescent Psychiatry*. 2003; 42(4):458–467. [PubMed: 12649633]
- Constantino JN, Przybeck T, Friesen D, Todd RD. Reciprocal social behavior in children with and without pervasive developmental disorders. *Journal of Developmental and Behavioral Pediatrics*. 2000a; 21(1):2–11. [PubMed: 10706343]
- Constantino JN, Przybeck T, Friesen D, Todd RD. Reciprocal social behavior in children with and without pervasive developmental disorders. *Journal of Developmental and Behavioral Pediatrics*. 2000b; 21(1):2–11. [PubMed: 10706343]
- Constantino JN, Todd RD. Genetic structure of reciprocal social behavior. *American Journal of Psychiatry*. 2000; 157(12):2043–2045. [PubMed: 11097975]

- Constantino JN, Todd RD. Autistic traits in the general population: a twin study. *Archives of General Psychiatry*. 2003; 60(5):524–530.
- Fenson L, Pethick S, Renda C, Cox JL, Dale PS, Resnick JS. Short-form versions of the MacArthur Communicative Development Inventories. *Applied Psycholinguistics*. 2000; 21:95–116.
- Frazier TW, Ratliff KR, Gruber C, Zhang Y, Law PA, Constantino JN. Confirmatory factor analytic structure and measurement invariance of quantitative autistic traits measured by the Social Responsiveness Scale-2. *Autism*. 2014; 18(1):31–44. [PubMed: 24019124]
- Hallmayer J, Cleveland S, Torres A, Phillips J, Cohen B, Torigoe T, et al. Genetic heritability and shared environmental factors among twin pairs with autism. *Archives of General Psychiatry*. 2011; 68(11):1095–1102. [PubMed: 21727249]
- Le Couteur A, Bailey A, Goode S, Pickles A, Robertson S, Gottesman I, Rutter M. A broader phenotype of autism: the clinical spectrum in twins. *Journal of Child Psychology and Psychiatry and Allied Disciplines*. 1996; 37(7):785–801.
- Lievens F, De Corte W, Schollaert E. A closer look at the frame-of-reference effect in personality scale scores and validity. *Journal of Applied Psychology*. 2008; 93(2):268–279. [PubMed: 18361631]
- Lyall K, Constantino JN, Weisskopf MG, Roberts AL, Ascherio A, Santangelo SL. Parental social responsiveness and risk of autism spectrum disorder in offspring. *Journal of the American Medical Association*. IN PRESS.
- Missouri QuickFacts from the US Census Bureau. Retrieved January. 2013; 22:2014. from <http://quickfacts.census.gov/qfd/states/29000.html>.
- Ozonoff S, Iosif A-M, Baguio F, Cook IC, Hill MM, Hutman T, Rogers SJ, Rozga A, Sangha S, Sigman M, Steinfeld MB, Young GS. A Prospective Study of the Emergence of Early Behavioral Signs of Autism. *Journal of the American Academy of Child & Adolescent Psychiatry*. 2010; 49(3):256–266. e252. [PubMed: 20410715]
- Pandey J, Verbalis A, Robins DL, Boorstein H, Klin AM, Babitz T, Chawarska K, Volkmar F, Green J, Barton M, Fein D. Screening for autism in older and younger toddlers with the Modified Checklist for Autism in Toddlers. *Autism*. 2008; 12(5):513535.
- Pierce K, Carter C, Weinfeld M, Desmond J, Hazin R, Bjork R, Gallagher N. Detecting, Studying, and Treating Autism Early: The One-Year Well-Baby Check-Up Approach. *The Journal of Pediatrics*. 2011; 159(3):458–465. e456. [PubMed: 21524759]
- Piven J, Palmer P, Jacobi D, Childress D, Arndt S. Broader autism phenotype: evidence from a family history study of multiple-incidence autism families. *American Journal of Psychiatry*. 1997; 154(2): 185–190. [PubMed: 9016266]
- Price TS, Freeman B, Craig I, Petrill SA, Ebersole L, Plomin R. Infant zygosity can be assigned by parental report questionnaire data. *Twin Research*. 2000; 3(3):129–133. [PubMed: 11035484]
- Robins DL, Fein D, Barton ML, Green JA. The Modified Checklist for Autism in Toddlers: an initial study investigating the early detection of autism and pervasive developmental disorders. *Journal of Autism and Developmental Disorders*. 2001; 31(2):131144.
- Ronald A, Happe F, Bolton P, Butcher LM, Price TS, Wheelwright S, Baron-Cohen S, Plomin R. Genetic heterogeneity between the three components of the autism spectrum: a twin study. *Journal of the American Academy of Child and Adolescent Psychiatry*. 2006; 45(6):691–699.
- Skuse DH, Mandy W, Steer C, Miller LL, Goodman R, Lawrence K, Emond A, Golding J. Social communication competence and functional adaptation in a general population of children: preliminary evidence for sex-by-verbal IQ differential risk. *Journal of the American Academy of Child and Adolescent Psychiatry*. 2009; 48(2):128–137. [PubMed: 19106766]
- Virkud YV, Todd RD, Abbacchi AM, Zhang Y, Constantino JN. Familial aggregation of quantitative autistic traits in multiplex versus simplex autism. *American Journal of Medical Genetics B: Neuropsychiatry Genetics*. 2009; 150B(3):328–334. 2819431.
- Werner E, Dawson G. Validation of the phenomenon of autistic regression using home videotapes. *Archives of General Psychiatry*. 2005; 62(8):889–895. [PubMed: 16061766]
- Wetherby AM, Brosnan-Maddox S, Peace V, Newton L. Validation of the Infant-Toddler Checklist as a broadband screener for autism spectrum disorders from 9 to 24 months of age. *Autism*. 2008; 12(5):487–511. [PubMed: 18805944]

Wolff S, Narayan S, Moyes B. Personality characteristics of parents of autistic children: a controlled study. *Journal of Child Psychology and Psychiatry and Allied Disciplines*. 1988; 29(2):143–153.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Key Points

- Reciprocal social behavior (RSB) is a developmental domain that emerges early in childhood and is disrupted in autism spectrum disorder (ASD).
- Video-referenced ratings of RSB offer a novel, rapid methodology for serial measurement of quantitative variation in RSB between 18 and 24 months.
- The capacity for RSB, as measured by this method, exhibits a continuous distribution, sex differences, and preliminary evidence suggesting heritability in toddlers, with higher scores characterizing toddlers at elevated risk of ASD.
- Implementation of a video-referenced rating system represents a promising, feasible approach to track the early development of RSB over time and in response to early intervention.



Figure 1. Video-referenced Standard for vrRSB

Still frames from the video portion of the vrRSB displaying the child serving as a frame of reference for developmentally appropriate social engagement. Key behaviors include, from left to right, turn-taking while throwing a ball, interactive symbolic play via feeding a doll, and sharing positive affect.

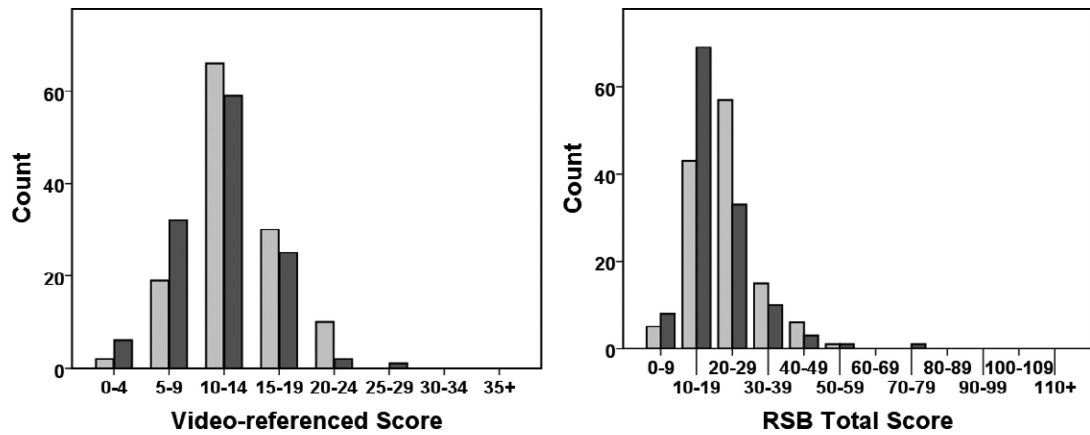


Figure 2. 18-Month vrRSB Scores Are Continuously Distributed

Scores for males and females are displayed for the 13 video-referenced items (1A) and the 44 questions in the full vrRSB (1B). The continuous distributions reflect the range of RSB present in toddlers. Higher scores indicate less RSB, and in this epidemiologic sample, there is a preponderance of lower scores. Males have on average significantly higher scores than females (video-referenced: 13.5 (4.0) vs. 11.4 (4.2); RSB Total: 24.1 (9.5) vs. 18.9 (8.1)) consistent with measurements of RSB in later childhood and adulthood. One high-scoring outlier is present in both panels, representing the same individual, whose scores on both the M-CHAT and MCDI-s indicated social and language delays, respectively.

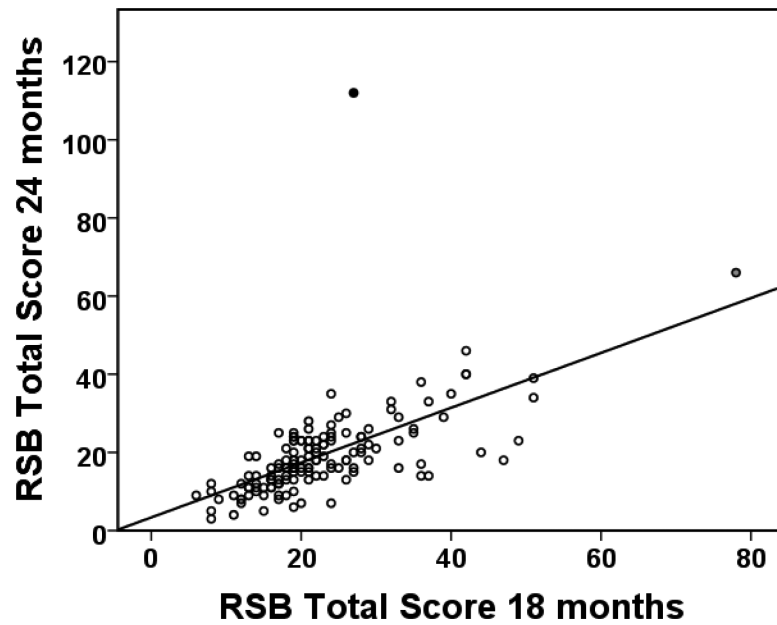


Figure 3. Scatterplot of vrRSB scores at 18 and 24 months

ICC=.704 ($p<.000$) for this twin sample following removal of two outliers (gray and black dots), consistent with excellent test-retest reliability. The overall sample shows a slight decrease in scores from 18 to 24 months, suggesting improvement attributable to developmental maturation. The high-scoring outlier observed in Figure 1 at 18 months (gray dot) also shows a high score at 24 months, whereas another outlier (black dot) demonstrated a dramatic increase in score between 18 and 24 months. Scores on the MCDI-s also worsened for this participant during this interval, dropping from typical levels to below the 1st percentile.

Table 1

Participant Characteristics

	%	n (twins)
Gender		
Male	50.4	127
Female	49.6	125
Zygosity		
Monozygous (31 pairs)	24.6	62
Dizygous (95 pairs)	75.4	190
Same Sex (58 pairs)	46.0	116
Opposite Sex (37 pairs)	29.4	74
Income		
\$29,999	11.3	28
\$30,000- \$59,999	24.2	60
\$60,000- \$89,999	27.4	68
\$90,000-\$119,999	21.8	54
\$120,000- \$149,999	7.3	18
\$150,000 - \$179,999	2.4	6
\$180,000- \$209,999	3.2	8
\$210,000- \$249,999	1.6	4
>\$250,000	0.8	2
Race		
Caucasian	82.9	209
Black/African-American	8.7	22
Asian	1.6	4
Ethnicity		
Non-Hispanic	93.7	236
Hispanic	5.6	14
	Mean (SD)	Range
Age at baseline (months)	18.7 (1.1)	17.4-22.6
Total RSB score	21.5 (9.3)	6.00-78.00
Video-referenced total score	12.5 (4.2)	0.00-29.00
M-CHAT Total Score	1.02 (1.60)	0-16
MCDI-s percentile words understood	43.99 (36.0)	0-99
MCDI-s percentile words produced	39.5 (27.6)	0-99

The proportions of male to female twin pairs do not differ compared to Missouri twins born between September and December 2011 and August and December of 2012, the source populations for the study ($\chi^2=.060$, $df=1$, $p=0.807$). The median household income was between \$60,000 and \$89,999, which is higher than the Missouri median of \$47,202. The racial breakdown in the study differs significantly from the Missouri population ($\chi^2=24.952$, $df=2$, $p=0.000$), with a greater percentage of African American children (22.3% vs. 8.2%). The ethnic composition does not significantly differ ($\chi^2=0.036$, $df=1$, $p=0.849$). RSB = reciprocal social behavior. M-CHAT = Modified Checklist for Autism in Toddlers. MCDI-s = MacArthur Communicative Development Infant Short Form.

Table 2

Regression Model for 18 month vrRSB Score

Variable	Video-referenced				RSB Total			
	Parameter Estimate	Standard Error	T	Significance	Parameter Estimate	Standard Error	T	Significance
Intercept	13.206	.950	13.89	<.0001	24.059	2.108	11.41	<.0001
Gender	2.319	.704	3.29	.0013	5.532	1.562	3.54	.0006
Ethnicity	.417	1.517	.27	.784	1.078	3.365	.32	.7492
Maternal Education-mid	.615	.839	.73	.465	-2.356	1.860	-1.27	.2078
Maternal Education-high	-.842	.938	-.90	.3714	-3.476	2.080	-1.67	.0974
Expressive Vocabulary	-.0493	.0129	-3.81	.0002	-.0864	.0287	-3.01	.0032

For video referenced scores, $R^2=.1951$, $p<.0001$; for RSB Total, $R^2=.1657$, $p<.0006$. Variance in gender and expressive language significantly contribute to both types of scores.

Table 3

Differences in Video-referenced and Total RSB Scores between 18 and 24 Months

	Time 1 Mean (SD)	Time 2 Mean (SD)	Statistics
One twin per family			
Total RSB Score	22.4 (8.4)	18.6 (7.6)	t=5.036, df=68, p<.000
	22.5 (9.6)	18.3 (8.5)	t=5.086, df=70, p<.000
Video-referenced RSB Score	13.3 (3.6)	10.7 (4.3)	t=5.511, df=68, p<.000
	13.0 (3.7)	9.9 (4.3)	t=5.858, df=70, p<.000
Social Communication Items	20.9 (6.8)	17.1 (6.8)	t=5.886, df=68, p<.000
	20.7 (7.8)	16.6 (7.3)	t=5.284, df=70, p<.000
Restricted Repetitive Items	1.5 (2.6)	1.4 (1.8)	t=-.218, df=68, p<.828
	1.7 (3.2)	1.6 (2.2)	t=.388, df=70, p<.699
One male twin per family			
Total RSB Score	23.5 (7.4)	19.8 (7.9)	t=4.133, df=42, p<.000
	23.4 (9.7)	19.3 (8.6)	t=4.096, df=43, p<.000
Video-referenced RSB Score	13.7 (4.0)	11.3 (4.5)	t=4.094, df=42, p<.000
	13.2 (3.8)	10.5 (4.4)	t=4.418, df=43, p<.000
Social Communication Items	22.1 (6.6)	18.8 (7.1)	t=4.798, df=42, p<.000
	21.5 (7.6)	17.5 (7.4)	t=4.568, df=43, p<.000
Restricted Repetitive Items	1.5 (2.0)	1.7 (2.0)	t=-.697, df=42, p<.489
	1.9 (3.8)	1.7 (2.3)	t=.415, df=43, p<.680
One female twin per family			
Total RSB Score	20.5 (9.8)	16.4 (6.8)	t=2.912, df=25, p<.007
	20.9 (9.4)	16.6 (8.0)	t=2.988, df=26, p<.006
Video-referenced RSB Score	12.4 (2.9)	9.7 (3.7)	t=3.664, df=25, p<.001
	12.7 (3.6)	8.9 (3.9)	t=3.854, df=26, p<.001
Social Communication Items	19.1 (6.9)	15.4 (6.0)	t=3.362, df=25, p<.002
	19.4 (8.1)	15.1(7.0)	t=2.898, df=26, p<.008
Restricted Repetitive Items	1.46 (3.4)	1.0 (1.2)	t=.840, df=25, p<.409
	1.5 (1.9)	1.5 (1.9)	t=.000, df=26, p<.1.00

Within cells, the first and second lines contain twin 1 and twin 2 data, respectively. Outliers (see Figure 3) are excluded. Decreases in video-referenced, RSB Total Scores, and SC items over time suggest developmental improvement in the capacity for RSB. RSB=reciprocal social behavior.

Table 4

Twin-twin Intraclass Correlations for RSB in early childhood

Twin type	Number of twin pairs	Video-referenced RSB	RSB Total Score	Social Communication Items	Restrictive Repetitive Behavior Items
MZ All	31	.917 (p<.000)	.863 (p<.000)	.879 (p<.000)	.467 (p<.004)
DZ All	95	.419 (p<.000)	.231 (p<.012)	.281 (p<.003)	.316 (p<.001)
Male MZ	16	.863 (p<.000)	.879 (p<.000)	.878 (p<.000)	.598 (p<.006)
Male DZ	29	.356 (p<.027)	.287 (P<.062)	.288 (p<.061)	.522 (p<.002)
Female MZ	15	.960 (p<.000)	.779 (P<.001)	.833 (p<.000)	.394 (p<.066)
Female DZ	29	.660 (P<.000)	.366 (P<.023)	.460 (p<.005)	.168 (p<.188)
OS	37	.358(p<.014)	.093 (p<.289)	.168(p<.156)	.107 (p<.262)

Numbers in parentheses indicate pairs of twins. MZ = monozygotic, DZ= dizygotic, OS = opposite sex.