HHS Public Access

Author manuscript

J Autism Dev Disord. Author manuscript; available in PMC 2019 June 01.

Published in final edited form as:

J Autism Dev Disord. 2018 June ; 48(6): 1932–1944. doi:10.1007/s10803-017-3458-9.

An Initial Psychometric Evaluation of the Joint Attention Protocol

Sallie W. Nowell[1], **Linda R. Watson**[1], **Richard A. Faldowski**[1], and **Grace T. Baranek**[2]

[1]The University of North Carolina at Chapel Hill, Department of Allied Health Sciences, Chapel Hill, NC, USA

[2]The University of Southern California, Chan Division of Occupational Science and Occupational Therapy, Los Angeles, CA, USA

Abstract

The goal of this paper is to examine the psychometric properties of a live-coded behavioral measure of joint attention, the *Attention-Following and Initiating Joint Attention Protocol* (JA Protocol), in order to assist researchers and clinicians in identifying when this measure may meet their joint attention assessment needs. Data from 260 children with autism spectrum disorder, developmental delay, or typical development between the ages of 2 and 12 years were used to evaluate this measure using quality standards for measurement. Overall, the JA Protocol demonstrated good psychometric properties. Recommendations and limitations for use of this measure based on psychometric analysis results are reported.

Keywords

joint attention; psychometric; measurement; autism; ASD

Introduction

Joint attention skills have been identified as pivotal in the development of all children, including those with autism spectrum disorder (ASD), and have thus become key early intervention targets for children with ASD. Knowledge about intervention effects in populations like those with ASD is dependent upon the measures from which results of

Correspondence concerning this article should be addressed to Sallie Nowell, Division of Speech & Hearing Sciences, CB# 7190, The University of North Carolina at Chapel Hill, Chapel Hill, NC, 27599-7190, nowell@unc.edu.
Sallie W. Nowell, The University of North Carolina at Chapel Hill, Chapel Hill, NC, USA
Linda R. Watson, The University of North Carolina at Chapel Hill, Chapel Hill, NC, USA
Richard A. Faldowski, The University of North Carolina at Chapel Hill, Chapel Hill, NC, USA
Grace T. Baranek, The University of Southern California, Los Angeles, CA, USA
Contact: nowell@unc.edu

Conflicts of Interest: Dr. Watson and Dr. Baranek have intellectual property rights related to the JA Protocol as authors. Ms. Nowell and Dr. Faldowski declare no conflicts of interest. The authors declare they have no financial conflicts of interest.

Compliance with Ethical Standards

Ethical approval: All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent: Informed consent was obtained from all individual participants included in the study.

research studies are drawn (Naglieri & Chambers, 2009); therefore, having measures that are psychometrically sound for the population of individuals with ASD is necessary. Furthermore, ASD is no longer considered a rare disorder that can only be diagnosed and treated at university clinics; thus, clinicians in community practice settings need access to reliable and valid measures for this population (Klin, Saulnier, Tsatanis, & Volkmar, 2005). Many observational measures of joint attention require extensive clinician training or time-intensive video coding that may be impractical for the settings in which interventions are taking place. The purpose of this article is to examine the psychometric properties of a live-coded behavioral measure of joint attention, the *Attention-Following and Initiating Joint Attention Protocol* (JA Protocol; BLINDED), using quality standards for measurement to assist researchers and clinicians in identifying when this measure may meet their joint attention assessment needs.

Joint Attention, A Pivotal Skill in ASD

Joint attention is a set of nonverbal behaviors including eye gaze, pointing, and showing, which are used to reference outside objects during a communicative exchange (Bruinsma, Koegel, & Koegel, 2004; Carpenter, Nagell, & Tomasello, 1998; Charman, 2003). Joint attention can be divided into acts of response and initiation (Brunisma et al., 2004). Initiation of joint attention (IJA) is when a child uses nonverbal means such as gestures and eye gaze to draw the attention of a communication partner to an object to share interest (Mundy & Newell, 2007). Response to joint attention (RJA) involves a child shifting their attention appropriately in response to another person's gestures and eye gaze (Bruinsma et al., 2004). Joint attention refers specifically to the cognitive skill of sharing attention, which requires understanding that a communication partner behaves intentionally as well as awareness of instances when they are sharing attention with that communication partner. Tomasello (1995) cautions that true joint attention behaviors are distinct from early "simultaneous looking" behaviors, typical of 6–9 month infants, wherein the child may look at the same stimulus in the environment as their parent (e.g. an airplane overhead or a toy being played with) and even shift visual attention between the parent and stimulus, but the child is yet not cognitively aware that the adult is sharing that experience or looks at things intentionally.

IJA and RJA behaviors are among the earliest intentional forms of communication that a child develops (Bruner, 1981). Early joint attention behaviors emerge during the last quarter of the first year in typical development. Between 8 and 9 months, infants begin using their eye gaze to direct the attention of caregivers and to demonstrate joint engagement with a caregiver and an object (Carpenter, Nagel, & Tomasello, 1998). Distal deictic gestures, which involve intentionally drawing another person's attention to an object out of reach such as pointing and reaching develop between 10 and 12 months and involve objects farther away from the child by 13–14 months (Masur, 1983; Carpenter et al, 1998; Crais, 2004; Crais, Watson, and Baranek, 2009). Carpenter and colleagues (1998) found that attention following behaviors emerge around 11 months of age with point following usually developing slightly before eye gaze following.

Based on the general developmental theory, the ability to coordinate attention between a person and an object by understanding and using nonverbal communication is a critical precursor to later language skills (Bates, 1979). Moving from a dyadic to a triadic interaction may also mark an early awareness of theory of mind (Baron-Cohen, 1991; Charman, 2003). In children with autism spectrum disorder (ASD), joint attention has been identified as a pivotal skill in that delayed or atypical development of joint attention is characteristic of the disorder and these delays have cascading effects on later social communicative abilities (Bruinsma et al., 2004; Charman, 2003; Poon, Watson, Baranek, & Poe, 2012). Research has established that early development of joint attention predicts later language and social cognitive skills in children with ASD (Charman, 2003; Mundy, Sigman, & Kasari, 1990; Sigman et al., 1999; Stone & Yoder, 2001); therefore, joint attention skills are critical targets of intervention for this population (Kasari, Gulsrud, Freeman, Paparella, & Heelemann, 2012) that require regular clinical assessment to monitor progress toward treatment goals.

Assessment of Children with Autism Spectrum Disorder

Some unique qualities of the ASD population should be considered during assessment selection and measure development (Klin et al., 2005). First, unlike their typically developing and developmentally delayed peers of unknown etiology, individuals with ASD tend to have uneven patterns of developmental strengths and needs throughout their lives (Ozonoff, Goodlin-Jones, & Solomon, 2010) that can complicate interpretation of composite scores on tests designed to summarize more than one skill area (Kasari, Brady, Lord, & Tager-Flusberg, 2013). Secondly, symptoms of ASD change over time, often independent of other developmental changes, making compliance with testing quite variable within individuals (Klin et al., 2005; Ozonoff et al., 2010).

Another important consideration is that the observational assessments available either involve extensive clinician training (e.g. Autism Diagnostic Observation Schedule, ADOS; Lord et al., 2000, 2002) or time consuming video coding (e.g. Early Social Communication Scales, ESCS; Mundy et al., 2003), which may not be feasible or practical in all clinical settings. Since joint attention development is delayed in children with ASD (Poon et al., 2012), it is also critical that preverbal communication assessments be validated for older preschool and school-aged children (Kasari et al., 2013). This study examines the psychometric properties of the JA Protocol, a measure of RJA and IJA that is live-coded for ease of use in clinical settings and may be valid for use with older children who have developmental delays.

Quality standards for clinical tools in the ASD population

Authors in disciplines such as speech-language pathology (Friberg, 2010; McCauley & Swisher, 1984; Plante & Vance, 1994), healthcare (DeVon et al., 2007; Kottner et al., 2011), and psychology (Cicchetti, 1994; Furr, 2011a) have proposed guidelines for clinical evaluation of measurement quality. These quality standards overlap substantially; therefore, we compiled guidelines for measurement evaluation from these sources (See Table 2).

The Attention-Following and Initiating Joint Attention Protocol

Based on previously studied joint attention tasks and assessment tools (Brady, Steeples, & Fleming 2005; Leekam, Hunnisett, & Moore, 1998; Mundy et al., 2003; Stone, Ousley, Yoder, Hogan, & Hepburn, 1997), BLINDED designed *The Attention-Following and Initiating Joint Attention Protocol* (JA Protocol) to measure a child's response to and initiation of joint attention bids during a semi-structured interaction with an examiner. The protocol alternates between eight initiation of joint attention (IJA) trials and eight attention-following (i.e. response to joint attention; RJA) trials, for a total of 16 opportunities to observe joint attention skills. RJA items gradually increase in level of support and prompting (e.g., pointing in addition to gaze shift) across the assessment. The JA Protocol was designed with four goals: 1) to specifically measure RJA and IJA, 2) to allow efficient live scoring during the administration instead of coding from video, 3) to provide a comfortable testing context with materials appropriate for a wide developmental range of children, and 4) to be sensitive to differences within and between diagnostic groups of children (BLINDED; BLINDED).

Step-by-step testing procedures with pictures and operationally defined item prompts and responses can be found in the JA Protocol manual (see BLINDED; web address BLINDED). A summary of the items can be viewed in Table 1. All items are scored dichotomously: "1" for pass or "0" for fail. The protocol is designed for live scoring by the examiner for ease of clinical use. It is intended for individual administration to a child in a separate room that can be set up for the assessment with novel visual stimuli. The child is seated at a table, with the examiner seated directly across from the child at approximately the same height. The typical administration time is 20 to 25 minutes.

The JA Protocol has been previously used as a measure of IJA and RJA in studies of children with ASD and other developmental disabilities (BLINDED; BLINDED). However, the psychometric properties of the JA Protocol were not fully examined or reported in these earlier studies. The aim of the present study was to evaluate the JA Protocol following the guidelines for quality standards summarized in Table 2. We hypothesized that: (1) The JA Protocol would meet Cicchetti and Sparrow's (1981, 1990) criteria for test-retest and interrater reliability at the item and subscale levels; (2) The JA Protocol subscale scores would demonstrate moderate to strong correlations with language and social communication measures to establish convergent validity and weak correlations with ADOS Restricted and Repetitive Behavior domain scores to establish divergent validity; (3) The dimensionality of the JA Protocol would be confirmed with a confirmatory factor analysis demonstrating the fit of a 2-factor model with RJA and IJA items loading on 2 different factors.

Method

The psychometric properties of the JA Protocol were examined using data collected from two samples, as detailed in subsequent sections. Analyses were conducted using JMP Pro 13.0 (2014) and Mplus 7 (Version 7.4; Muthen and Muthen, 2015).

Sample 1

Data for Sample 1 were collected as part of a larger assessment protocol for (BLINDED). The order of assessments was randomized over the course of 1–2 clinic visits and the JA Protocol was completed during one of those clinic visits. A pool of 11 coders collected data for Sample 1. The majority of assessments were administered and coded by research assistants with bachelor's degrees who had experience working with individuals with ASD and achieved reliability on the JA Protocol. In some cases, master's students or doctoral-level research assistants in an autism-related field (i.e. occupational therapy, special education, or speech-language pathology) administered and coded assessments. The reliability process for Sample 1 was as follows: first, the coders reviewed the drafted manual and discussed questions with the test developer. Since the measure was originally developed for use in the study from which Sample 1 was derived, the manual was still being developed and no "gold standard" scoring criterion were established at the start of the study. Thus, the first few recordings of the JA Protocol were consensus scored by the coding team and test developer as part of the training and manualization process. For subsequent coders who joined the team, reliability was achieved by agreement on ≥80% of the item scores (at least 13 of 16 item scores) of two sequential tapes from the consensus scored assessment recordings. Administration training consisted only of reviewing the manual, watching the videos for the scoring reliability process, and practicing an administration alone or with a partner using the testing materials. The assessments were video recorded and independently scored by a second coder at a later date. Coders were not naive the child's diagnosis.

Sample 1 consisted of 260 children from [BLINDED] who were either typically developing (n = 81), developmentally delayed (n = 57), or diagnosed with ASD (n = 122). To be included in the parent study, the children with ASD had confirmed diagnoses based on the Autism Diagnostic Observation Schedule (ADOS; Lord et al., 2000; 2012) and the Autism Diagnostic Interview- Revised (ADI-R; Lord et. al., 1994). The sample of children with ASD had mean scores on developmental and language assessments indicative of severe levels of general developmental and language delays; none of the children with ASD had language skills beyond an ADOS module 2 "phase speech" level. The DD group consisted of children with known genetic syndromes (e.g. Down syndrome, Prader-Willi syndrome, Williams syndrome) and developmental disorders as well as children with idiopathic developmental delay (i.e. >2 SD below the mean on an IQ measure, or >1.5 SD below the mean in 2 or more developmental domains). Children in the DD group did not meet criteria for an ASD diagnosis based on an assessment using the ADI and ADOS. To be included in the TD group, children must neither have been receiving services through their school system nor had any significant developmental problems diagnosed.

Demographic data for Sample 1 are summarized in Table 3. Ages ranged from 10 months to 12 years. The majority of the children across diagnostic groups were white (80%), had mothers who pursued further education after high school (>70%), and had a household income ≥$60,000 (>55%). No significant group differences were found using one-way ANOVAs on demographic variables that were thought to affect test score analyses. The Preschool Language Scales 4th Edition (PLS-4; Zimmerman, Steiner, & Pond, 2002) as well as language subscales of the Mullen Scales of Early Learning (Mullen, 1995) and Vineland

Adaptive Behavior Scales (VABS; Sparrow, Balla, & Cicchetti, 1984) were administered to all children in this sample (see Table 4 for a summary of Sample 1 scores on these measures). Data from Sample 1 were used to examine inter-rater reliability, internal consistency, convergent and divergent validity, and dimensionality of the JA Protocol.

Sample 2

Sample 2 consisted of 19 children from central [BLINDED] with clinical diagnoses of ASD (Sample 2 Demographics in Table 5). All of the children in this sample were male and ranged in age from 3 to 7 years (M = 71.21 months). They were recruited through an autism research registry specifically for the purpose of examining the test-retest reliability of the JA Protocol. The exclusion of females in this sample was not intentional; no parents of female children responded to our research registry request for study participants. The only additional measure gathered on children in Sample 2 was the Social Communication Questionnaire (SCQ; Rutter et al., 2003), which was completed by parents at the child's first of the two clinic visits during which the JA Protocol was administered.

Children completed the JA Protocol twice with approximately two weeks between assessments. The two-week time interval was chosen to give children time to forget the first administration while not giving children enough time to develop more joint attention skills than they had at the first administration of the protocol. Two examiners were present in the room during the assessment. The primary examiner administered the JA Protocol at both of the child's visits and scored it live while the secondary examiner observed the administration of the JA Protocol and independently scored it live at both visits to collect interrater reliability data.

Examiners and coders for Sample 2 were 3 female Master's students in speech-language pathology with experience working with children with ASD. Each of the examiners was trained to administer and score the JA Protocol by the test developer and second author; they established scoring reliability (≥80% agreement; i.e. agreement on at least 13 of the 16 item scores) on at least two sequential training DVDs of children recorded from Sample 1 before collecting data. Scoring training consisted of reviewing the JA Protocol manual and discussing questions with the test developer before watching and scoring reliability recordings. The master's students watched between 2 and 5 recordings before achieving scoring reliability. Coders were aware that children in the sample had diagnoses of ASD.

Measures

The *Autism Diagnostic Observation Schedule* (ADOS, Lord et al., 2000, 2002) is a semi-structured observational measure of symptoms of ASD. Items from measure were used to determine convergent and divergent validity of the JA Protocol. Modules of the ADOS are administered according to a child's language level. The ADOS Module 1 ("no words" or "single words" language level) or 2 ("phrase speech" language level) was administered to the children with ASD and DD in this study according to their language level. For convergent validity analyses, the following ADOS joint attention items were used: *Pointing, Showing, Response to Joint Attention,* and *Spontaneous Initiation of Joint Attention*. Module 1 and 2 scores on these items were combined across the sample into a single item

variable to fully capture sample variability. A negative correlation was expected for the relationship between ADOS items and JA Protocol subscale scores since higher ADOS item scores indicate worse JA performance, while higher JA Protocol scores indicate better JA performance.

Items from the ADOS Restricted and Repetitive Behavior (RRB) domain were correlated with the JA Protocol IJA and RJA means to determine divergent validity. These items were selected because there were no theoretical reasons to hypothesize that JA Protocol scores would be strongly correlated with repetitive behaviors, and the ADOS RRB item scores could be directly compared with the ADOS Social Affective item scores used for convergent validity. Item scores across ADOS modules 1 and 2 were combined to capture the full sample variability. We also calculated the ADOS RRB calibrated severity score using the guidelines in Hus et al. (2014) to examine divergent validity. The ADOS RRB calibrated severity score standardizes the total RRB score to a 1–10 scale, which can be compared across chronological ages and ADOS modules as a measure of RRB symptom severity (i.e. the higher the score, the more severe the symptoms). Thus, the RRB calibrated score was used because it is not intended to be related to the child's developmental level, language level, or to JA specifically. In contrast, the JA Protocol is intended to be sensitive to developmental changes in the construct with higher scores indicating better JA skill performance so JA Protocol scores were expected to have a weak negative correlation with the RRB calibrated severity score from the ADOS. We expected that item level correlations would be stronger than the calibrated severity score since it was derived from multiple items.

The *Mullen Scales of Early Learning* (MSEL; Mullen, 1995) is an assessment of early childhood development in 5 domains: Visual Reception, Fine Motor, Receptive Language, Expressive Language, and Gross Motor. An Early Learning Composite (ELC) can be derived from the subscales (excluding Gross Motor) as an overall developmental skills estimate. The MSEL was administered to all children in the study. The raw scores from the Receptive and Expressive Language domains of the MSEL were used as measures of convergent validity in the present study analyses.

The *Preschool Language Scales- Fourth Edition* (PLS-4; Zimmerman, Steiner, and Pond, 2002) is a measure of receptive and expressive language skills, which was administered to children in all study groups who were within the 2–6 year old age range for which this measure is validated. It includes 2 domain scores: Auditory Comprehension and Expressive Communication. Raw scores from these domains were used in convergent validity analyses in the study analyses.

The *Social Communication Questionnaire* (SCQ; Rutter et al., 2003) is a 40-item parent report screening for symptoms of ASD. Scores above the cutoff of 15 indicate that an individual is likely to have ASD. The SCQ was completed by parents of children in Sample 2 at their first clinic to measure ASD symptom severity in the participants.

Results

Reliability

Internal consistency—Cronbach's alpha for Sample 1 was .91, meeting Cicchetti and Sparrow's (1990) psychometric criteria for "excellent" statistical significance. Internal consistency was also calculated by subscale: the IJA subscale was in the "good" range (Cronbach's alpha = .86) while the RJA subscale was in the "excellent" range (Cronbach's alpha = .90).

Test-retest reliability—These data were only collected from Sample 2 with a test-retest interval of approximately two weeks. ICCs were calculated for the IJA and RJA subscale means as well as individual items (see Table 6). The ICCs for the RJA (.80) and IJA (.83) subscales were in the "good" range of agreement according to Cicchetti and Sparrow (1981).

Interrater reliability—Slightly more than 20% of the assessments for Sample 1 were quasi-randomly selected (i.e. one of approximately every 5 were selected) to be re-coded from video for interrater reliability (n = 56). ICCs were calculated for Sample 1 items and subscale means (see Table 7). Subscale (RJA = .97; IJA = .89) and item level ICCs were all in the "good" to "excellent" range except for the first IJA item (.51), which was in the "fair" range. These ICCs, which assume that raters were paired more systematically than they were in reality, may underestimate the true interrater reliability of the sample. For Sample 2, interrater reliability data was collected during test and retest protocol administrations for each of the 19 participants. Subscale reliability ICCs were .99 for RJA and .98 for IJA.

Validity

Convergent Validity—Convergent validity was calculated for Sample 1 by correlating the mean RJA and IJA subscale scores with Sample 1 children's scores on other measures of language skills, IJA, and RJA (Table 8). Negative correlations are expected between ADOS item scores and IJA and RJA subscale scores of the JA Protocol since higher ADOS scores indicate greater symptom severity while higher JA Protocol subscale scores indicate better use of JA. *Response to Joint Attention* on the ADOS was strongly correlated with the JA Protocol RJA mean for the full (i.e. ASD and DD combined) sample (r = −.69) and even more strongly for the subsample of children diagnosed with ASD (r = −.80) as compared to the DD subsample (r = −.37). Furthermore, the *Spontaneous Initiation of Joint Attention* on the ADOS was moderately correlated with the IJA mean for the ASD and DD combined sample (r = −.50), but again, more strongly correlated with the ASD subsample (r = −.39) than the DD subsample (r = −.05), indicating good convergent validity.

Because joint attention has demonstrated strong concurrent and predictive relationships with child language skills in the literature, we hypothesized that the JA Protocol subscale means (i.e. IJA and RJA mean scores) would be at least moderately correlated with the PLS-4 raw scores and Mullen Scales of Early Learning raw scores for receptive and expressive language. These language measures demonstrated stronger associations with the RJA mean than the IJA mean, but both met hypothesized expectations of moderate to strong correlations with language measures. Notably, the associations between language skills and

JA Protocol subscales were strongest in the ASD subsample and decreased in magnitude for the DD and TD groups (Table 8). This finding may support the differential importance of the joint attention as a pivotal skill for language acquisition for children with ASD as compared to those in the DD and TD groups.

Divergent Validity—For the ASD subsample, the majority of the ADOS RRB scores were more weakly associated with IJA and RJA means than was true for the ADOS Social Affective and language measures (see Table 8). Item level scores, particularly for *Unusual Sensory Interest* and *Hand/Finger Mannerisms,* had weaker correlations with the JA Protocol items than the RRB Calibrated Severity Score. The magnitude of the ASD subsample correlations between the JA Protocol and the ADOS RRB were stronger than for the DD subsample.

Examination of the Dimensionality of the Measure

To examine the dimensionality of the JA Protocol, tetrachoric correlations between the binary items with each other and biserial correlations between the binary items and the subscale (IJA and RJA) means were explored with each sample. Tetrachoric correlations and biserial correlations are reported because, with binary items, Pearson correlations underestimate the true relationship between the binary items or between binary items and continuous IJA and RJA subscale scores (Bonett, 2005; Brown, 1977; Drasgow, 1988; Harris, 1988; Uebersax, 2017). As hypothesized, items from the IJA subscale were more strongly correlated with other IJA items and the IJA mean score than RJA items and vice versa for both samples (Sample 1, Table 9 for tetrachoric correlations).

Sample 1 data were analyzed using confirmatory factor analysis to determine how well the theoretical 2-factor model with RJA and IJA subscales fit the data. Mplus 7.4 was used for these analyses to account for the binary nature of the JA Protocol items. A likelihood ratio χ^2 test of a one-factor model versus a two-factor model, with IJA and RJA items loading on different factors, strongly suggested that the two-factor was preferred (χ^2 [1, n = 253] = 22.330, $p \leq .001$). Moreover, the two-factor model was strongly supported by goodness of fit indices. The root mean square of error of approximation (RMSEA) was .01 (“acceptable” cut-off ≤ .05). Both the comparative fit index (CFI) and the Tucker-Lewis index (TLI) had values of 1.0 (“acceptable” cutoffs for both ≥ .90). All factor loadings in this model were approximately 1 (all loading p-values ≤ .001) and the correlation between the IJA and RJA subscales was a moderate .51 ($p \leq .001$), supporting the construct validity of the JA Protocol with this sample (Table 10).

Conclusions

Overall, the JA Protocol demonstrated good psychometric properties and promise for the clinical utility of this measure. The correlations of the JA Protocol RJA and IJA subscales with ADOS items measuring joint attention skills, along with the factor analysis results, support its construct validity. In considering the magnitude of correlations of the JA Protocol subscales with the ADOS items, it is important to remember that the ADOS variables reflect single items. In general, adding more items to measure a given construct (e.g., 8 opportunities to observe both RJA and IJA in the JA Protocol) is expected to improve the

reliability of measurement. An important future research step will be to compare scores on the JA Protocol to scores on other comprehensive measures of JA such as the ESCS (Mundy, 2003) and the Communication and Symbolic Behavior Scales (CSBS; Wetherby and Prizant, 2002).

Even though a discriminant function analysis revealed that the JA Protocol discriminated children with ASD from children without ASD with 85% accuracy in toddlers and preschoolers and 74% accuracy using the whole sample of children up to age 12 years, the JA Protocol is not intended for use for diagnostic or educational placement purposes. Rather, it was developed as a relatively efficient tool to provide scores that would quantify the RJA and IJA performance of children from ages 2 to 12 years, particularly children with ASD. Thus, we have not attempted to ascertain cut-off scores to provide the best discrimination among diagnostic groups, or to evaluate the sensitivity, specificity, and positive and negative predictive values for the measure. Further, in evaluating the application of this measure to other samples, or for other purposes, it is important to recognize that our sample of children with ASD had mean scores on the PLS-4 and MSEL indicative of severe levels of general developmental and language delays. Therefore, the JA Protocol may not be as useful in the assessment of JA performance among children with ASD with typical or mildly delayed cognitive and language functioning as for children like those in the current samples. In addition, the JA Protocol has more limited utility for assessment of children without ASD in the 2 to 12 year age range; in particular, typically developing children show limited variability or growth in performance across the age range in our cross-sectional data from Sample 1, and children with other DD also have mean scores in the upper range of possible scores on the JA Protocol across the age range that we assessed.

Our use of the JA Protocol thus far has been for research purposes. For example, in a previously published study of children with ASD and those with other DD in Sample 1, we used the JA Protocol to study the association between sensory hypo-reactivity and JA performance, in models that also included mental and chronological age. Even after accounting for mental and chronological age, we found a significant negative association between hypo-reactivity and both RJA and IJA (BLINDED).

We have not yet evaluated the JA Protocol in a clinical practice setting, but hope to do so in future research. In considering potential clinical applications, the protocol may be helpful in documenting the need for treatment goals addressing RJA and/or IJA in children with suspected or confirmed ASD and potentially to measure progress toward those goals. Children with ASD in the current study demonstrated wide variability in their JA performance and this variability was moderately to strongly associated with other measures of their language and communication functioning. This suggests that the JA Protocol may have utility for measuring change over time, but we have no direct evidence documenting this. Thus, evaluating the sensitivity of the JA Protocol to change associated with interventions is a future research step.

The JA Protocol stimuli may also provide clinicians with an idea of what types of materials and activities engage the child’s interest, for use in targeting joint attention through

intervention. In addition, the RJA trial prompting levels offer valuable information about what level of support the child needs for responding to joint attention.

Presently there are some limitations for use and score interpretation of the JA Protocol. The pass/fail scoring of each item contributes to the feasibility and likely the reliability of scoring during administration, but this scoring does not provide information about the quality of joint attention skills demonstrated. To supplement the JA Protocol data with parent reported data about the frequency of JA behaviors demonstrated by the child in various environments, we recommend the Childhood Joint-Attention Rating Scale (C-JARS; Mundy et al., 2017). Furthermore, the psychometric samples used in this study were not nationally representative and included limited parental education, racial, and socioeconomic diversity. More psychometric information is needed to confirm that the JA Protocol is appropriate for use with culturally, geographically, and linguistically diverse populations.

Based on our experience with the JA Protocol and these findings, we recommend that the JA Protocol be administered to children one-on-one in a room free from distraction except for the protocol stimuli. It could be administered in isolation or as part of a larger speech-language or ASD assessment. Children should have typical or sufficiently augmented vision, hearing, and head control and be able to sit independently in a child-sized chair to complete the protocol. It may be most useful in assessing children who are not yet speaking in complete spontaneous sentences. The chronological age of the child can extend into middle school because the prompts and stimuli used for the JA Protocol can be adapted to meet an older child's interests. It is recommended that clinicians administering the JA Protocol have at least a clinical master's student education level and experience administering tests to children with developmental disabilities including ASD. Based on the psychometric quality standards examined, the psychometric properties of the JA Protocol are strong and the administration and scoring time required is clinically realistic for children who are candidates for JA interventions.

Acknowledgments

Funding: This research was supported in part by grants from The National Institute for Child Health and Human Development. R01-HD42168 to Grace Baranek and P30-HD003110 to the Carolina Institute for Developmental Disabilities (current PI: Gabriel Dichter) at the University of North Carolina at Chapel Hill.

This research was supported in part by grants from The National Institute for Child Health and Human Development (R01-HD42168, P30-HD003110) to the University of North Carolina at Chapel Hill. The authors would like to thank the families who participated in the Sensory Experiences Project. The authors also gratefully acknowledge the contributions of Rachel Rhyne in refining the Attention Following and Initiating Joint Attention Protocol manual and for her leadership in collecting test-retest data.

References

American Psychiatric Association. Diagnostic and statistical manual of mental disorders, fifth edition (DSM-5®). Washington, D.C: American Psychiatric Publishing; 2013.

Arnold LE, Aman MG, Martin A, Collier-Crespin A, Vitiello B, Tierney E, … Klin A. Assessment in multisite randomized clinical trials of patients with autistic disorder: the Autism RUPP Network. Journal of Autism and Developmental Disorders. 2000; 30(2):99–111. [PubMed: 10832774]

Baron-Cohen S. Precursors to a theory of mind: Understanding attention in others. 1991

Bates, E. Language and context. New York: Academic Press; 1976.

Bonett DG, Price RM. Inferential methods for the tetrachoric correlation coefficient. Journal of Educational and Behavioral Statistics. 2005; 30:213.

Brown MB. Algorithm AS 116: The tetrachoric correlation and its standard error. Applied Statistics. 1977; 26:343–351.

Bracken BA. Limitations of preschool instruments and standards for minimal levels of technical adequacy. Journal of Psychoeducational Assessment. 1987; 5(4):313–326.

Bruner J. The social context of language acquisition. Language & Communication. 1981; 1(2):155–178.

Bruinsma Y, Koegel RL, Koegel LK. Joint attention and children with autism: A review of the literature. Mental retardation and developmental disabilities research reviews. 2004; 10(3):169–175. [PubMed: 15611988]

Carpenter M, Nagell K, Tomasello M. Social cognition, joint attention, and communicative competence from 9 to 15 months of age. Monographs of the society for research in child development. 1998:i–174.

Charman T. Why is joint attention a pivotal skill in autism? Philosophical Transactions of the Royal Society of London B: Biological Sciences. 2003; 358(1430):315–324. [PubMed: 12639329]

Cicchetti DV. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. Psychological assessment. 1994; 6(4):284.

Cicchetti DV, Sparrow SA. Developing criteria for establishing interrater reliability of specific items: applications to assessment of adaptive behavior. American journal of mental deficiency. 1981

Cicchetti, DV., Sparrow, SS. Assessment of adaptive behavior in young children. Cicchetti, Domenic V., Sparrow, Sara S., editors. 1990.

Cohen J. A coefficient of agreement for nominal scales. Edu ond Psych Meas. 1960; 20:37–46.

Cohen J. Weighted kappa: Nominal scale agreement with provision for scaled disagreement or partial credit. Psych Bull. 1968; 70:213–220.

Crais E, Douglas DD, Campbell CC. The intersection of the development of gestures and intentionality. Journal of Speech, Language, and Hearing Research. 2004; 47(3):678–694.

Crais ER, Watson LR, Baranek GT. Use of gesture development in profiling children's prelinguistic communication skills. American Journal of Speech-Language Pathology. 2009; 18(1):95–108. [PubMed: 19029535]

Cronbach LJ. Coefficient alpha and the internal structure of tests. Psychometrika. 1951; 16(3):297–334.

Johnson, James H., Goldman, Jacquelin, editors. Developmental assessment in clinical child psychology: A handbook. Elmsford, NY, US: Pergamon Press; 1990. p. 173-196.p. xp. 289

Dawson G, Toth K, Abbott R, Osterling J, Munson J, Estes A, Liaw J. Early social attention impairments in autism: social orienting, joint attention, and attention to distress. Developmental psychology. 2004; 40(2):271. [PubMed: 14979766]

DeVon HA, Block ME, Moyle-Wright P, Ernst DM, Hayden SJ, Lazzara DJ, … Kostas-Polston E. A psychometric toolbox for testing validity and reliability. Journal of Nursing scholarship. 2007; 39(2):155–164. [PubMed: 17535316]

Drasgow, F. Polychoric and polyserial correlations. In: Kotz, L., Johnson, NL., editors. Encyclopedia of Statistical Sciences. Vol. 7. New York: Wiley; 1988. p. 69-74.

Edelen MO, Reeve BB. Applying item response theory (IRT) modeling to questionnaire development, evaluation, and refinement. Quality of Life Research. 2007; 16(1):5–18. [PubMed: 17375372]

Filipek PA, Accardo PJ, Ashwal S, Baranek GT, Cook EH, Dawson G, … Levy SE. Practice parameter: Screening and diagnosis of autism Report of the Quality Standards Subcommittee of the American Academy of Neurology and the Child Neurology Society. Neurology. 2000; 55(4): 468–479. [PubMed: 10953176]

Friberg JC. Considerations for test selection: How do validity and reliability impact diagnostic decisions? Child Language Teaching and Therapy. 2010; 26(1):77–92.

Furr, M. Scale construction and psychometrics for social and personality psychology. SAGE Publications Ltd; 2011a. Core principles, best practices, and an overview of scale construction; p. 4-16.

Furr, M. Scale construction and psychometrics for social and personality psychology. SAGE Publications Ltd; 2011b. Threats to psychometric quality; p. 67-77.

Gorsuch, RL. Factor analysis. 2. Hillsdale, NJ: LEA; 1983.

Hallgren KA. Computing inter-rater reliability for observational data: an overview and tutorial. Tutorials in quantitative methods for psychology. 2012; 8(1):23. [PubMed: 22833776]

Harris, B. Tetrachoric correlation coefficient. In: Kotz, L., Johnson, NL., editors. Encyclopedia of Statistical Sciences. Vol. 9. New York: Wiley; 1988. p. 223-225.

Hurwitz S, Watson LR. Joint attention revisited: Finding strengths among children with autism. Autism. 2015 1362361315593536.

Hus V, Gotham KC. Standardizing ADOS domain scores: Separating severity of social affect and restricted and repetitive behaviors. Journal of Autism and Developmental Disorders. 2014:1–13.

Kasari C, Brady N, Lord C, Tager-Flusberg H. Assessing the minimally verbal school-aged child with autism spectrum disorder. Autism Research. 2013; 6(6):479–493. [PubMed: 24353165]

Kasari C, Gulsrud A, Freeman S, Paparella T, Hellemann G. Longitudinal follow-up of children with autism receiving targeted interventions on joint attention and play. Journal of the American Academy of Child & Adolescent Psychiatry. 2012; 51(5):487–495. [PubMed: 22525955]

Klin A, Saulnier C, Tsatsanis K, Volkmar FR. Clinical evaluation in autism spectrum disorders: Psychological assessment within a transdisciplinary framework. Handbook of autism and pervasive developmental disorders. 2005; 2:772–798.

Kottner J, Audigé L, Brorson S, Donner A, Gajewski BJ, Hróbjartsson A, … Streiner DL. Guidelines for reporting reliability and agreement studies (GRRAS) were proposed. International journal of nursing studies. 2011; 48(6):661–671. [PubMed: 21514934]

Kuder GF, Richardson MW. The theory of the estimation of test reliability. Psychometrika. 1937; 2(3): 151–160.

Leekam SR, Hunnisett E, Moore C. Targets and Cues: Gaze-following in Children with Autism. Journal of Child Psychology and Psychiatry. 1998; 39(7):951–962. [PubMed: 9804028]

Lord C, Risi S, Lambrecht L, Cook EH, Leventhal BL, DiLavore PC, et al. The Autism Diagnostic Observation Schedule–Generic: A standard measure of social and communication deficits associated with the spectrum of autism. Journal of Autism and Developmental Disorders. 2000; 30:205–223. [PubMed: 11055457]

Lord, C., Rutter, M., DiLavore, PC., Risi, S. Autism Diagnostic Observation Schedule manual. Los Angeles: Western Psychological Services; 2002.

MacDonald R, Anderson J, Dube WV, Geckeler A, Green G, Holcomb W, … Sanchez J. Behavioral assessment of joint attention: A methodological report. Research in Developmental Disabilities. 2006; 27(2):138–150. [PubMed: 15979843]

Masur EF. Gestural development, dual-directional signaling, and the transition to words. Journal of Psycholinguistic Research. 1983; 12(2):93–109.

McCauley RJ, Swisher L. Psychometric review of language and articulation tests for preschool children. Journal of Speech and Hearing Disorders. 1984; 49(1):34–42. [PubMed: 6700200]

Meindl JN, Cannella-Malone HI. Initiating and responding to joint attention bids in children with autism: A review of the literature. Research in developmental disabilities. 2011; 32(5):1441–1454. [PubMed: 21450441]

Mullen, EM. Mullen Scales of Early Learning. Circle Pines, MN: American Guidance Service; 1995.

Mundy, P., Delgado, C., Block, J., Venezia, M., Hogan, A., Seibert, J. Early social communication scales (ESCS). Coral Gables, FL: University of Miami; 2003.

Mundy P, Newell L. Attention, joint attention, and social cognition. Current directions in psychological science. 2007; 16(5):269–274. [PubMed: 19343102]

Mundy P, Novotny S, Swain-Lerro L, McIntyre N, Zajic M, Oswald T. Joint-Attention and the Social Phenotype of School-Aged Children with ASD. Journal of autism and developmental disorders. 2017; 47(5):1423–1435. [PubMed: 28229349]

Mundy P, Sigman M, Kasari C. A longitudinal study of joint attention and language development in autistic children. Journal of autism and developmental disorders. 1990; 20(1):115–128. [PubMed: 2324051]

Naglieri JA, Chambers KM. Psychometric issues and current scales for assessing autism spectrum disorders. Assessment of autism spectrum disorders. 2009:55–90.

Nunnally, JC., Bernstein, IH., Berge, JMT. Psychometric theory. 3. New York: McGraw-Hill; 1994.

Ozonoff S, Goodlin-Jones BL, Solomon M. Evidence-based assessment of autism spectrum disorders in children and adolescents. Journal of Clinical Child and Adolescent Psychology. 2005; 34(3): 523–540. [PubMed: 16083393]

Plante E, Vance R. Selection of Preschool Language Tests A Data-Based Approach. Language, Speech, and Hearing Services in Schools. 1994; 25(1):15–24.

Poon KK, Watson LR, Baranek GT, Poe MD. To what extent do joint attention, imitation, and object play behaviors in infancy predict later communication and intellectual functioning in ASD? Journal of autism and developmental disorders. 2012; 42(6):1064–1074. [PubMed: 21858585]

Prizant BM, Wetherby AM. Communicative intent: A framework for understanding social-communicative behavior in autism. Journal of the American Academy of Child & Adolescent Psychiatry. 1987; 26(4):472–479. [PubMed: 3308831]

Reeve, BB. An introduction to modern measurement theory. National Cancer Institute; 2002.

Rutter, M., Bailey, A., Lord, C. Social Communication Questionnaire. Los Angeles, CA: Western Psychological Services; 2003.

Seibert JM, Hogan AE, Mundy PC. Assessing interactional competencies: The early social-communication scales. Infant Mental Health Journal. 1982; 3(4):244–258.

Sigman M, Ruskin E, Arbelle S, Corona R, Dissanayake C, Espinosa M, … Robinson BF. Continuity and change in the social competence of children with autism, Down syndrome, and developmental delays. Monographs of the society for research in child development. 1999:i–139.

Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. Psychological bulletin. 1979; 86(2):420. [PubMed: 18839484]

Sparrow, S., Balla, D., Cicchetti, D. Vineland Adaptive Behavior Scales. Circle Pines, MN: American Guidance Service; 1984.

Stone WL, Coonrod EE, Ousley OY. Brief report: screening tool for autism in two-year-olds (STAT): development and preliminary data. Journal of autism and developmental disorders. 2000; 30(6): 607–612. [PubMed: 11261472]

Stone WL, Yoder PJ. Predicting spoken language level in children with autism spectrum disorders. Autism. 2001; 5(4):341–361. [PubMed: 11777253]

Streiner DL. Being inconsistent about consistency: When coefficient alpha does and doesn't matter. Journal of personality assessment. 2003; 80(3):217–222. [PubMed: 12763696]

Tager-Flusberg H, Kasari C. Minimally verbal school-aged children with autism spectrum disorder: the neglected end of the spectrum. Autism Research. 2013; 6(6):468–478. DOI: 10.1002/aur.1329 [PubMed: 24124067]

Tomasello M. Joint attention as social cognition. Joint attention: Its origins and role in development. 1995:103130.

Thompson B. Guidelines for authors reporting score reliability estimates. Score reliability: Contemporary thinking on reliability issues. 2003:91–101.

Thompson B. Understanding reliability and coefficient alpha, really. Score reliability: Contemporary thinking on reliability issues. 2003:3–23.

Uebersax, JS. [Accessed on 7/5/2017] Introduction to the tetrachoric and polychoric correlation coefficients. 2017. http://www.john-uebersax.com/stat/tetra.htm

Volkmar F, Cook EH, Pomeroy J, Realmuto G, Tanguay P. Practice parameters for the assessment and treatment of children, adolescents, and adults with autism and other pervasive developmental disorders. Journal of the American Academy of Child & Adolescent Psychiatry. 1999; 38(12): 32S–54S. [PubMed: 10624084]

Warren, SF., Yoder, PJ. Facilitating the transition from preintentional to intentional communication. In: Wetherby, AM.Warren, SF., Reichle, J., editors. Transitions in prelinguistic communications. Baltimore, MD: Brookes; 1998. p. 365-384.

Wetherby, A., Prizant, B. Communication and symbolic behavior scales—Normed edition. Baltimore, MD: Paul H. Brookes; 1993.

Wetherby, A., Prizant, B. Communication and symbolic behavior scales developmental profile—First Normed edition. Baltimore, MD: Paul H. Brookes; 2002.

Zimmerman, IL., Steiner, VG., Pond, RE. Preschool Language Scale–Fourth Edition Examiner's Manual. San Antonio, TX: Psychological Corporation; 2002.

Table 1

Summary of Joint Attention Protocol Items

Item	Type	Description
1	IJA	*Toy in marker box*: Examiner and Child engage in coloring with markers, with Examiner modeling and encouraging lifting lid of box to get markers of different colors. Rater 30–60 seconds, Examiner slips decorated cylindrical toy into marker box and offers closed marker box to Child. If Child opens box and does not initiate JA, Examiner waits short time and then offers marker box again.
2	RJA	Examiner gets Child's attention (eye contact or visual orientation to Examiner). Then Examiner LOOKS ONLY at picture on wall to right.
3	IJA	*Ball drop game:* Examiner offers balls to Child 1 by 1 to put down ball drop chute. Rater 3rd time, Examiner offers Child a ball too large for chute. If Child does not initiate JA, sequence is repeated 2nd time.
4	RJA	Examiner gets child's attention (eye contact or visual orientation to Examiner). Then Examiner LOOKS ONLY at a picture on wall to the left.
5	IJA	*Dress up:* Examiner opens box filled with a variety of scarves, hats, gloves, glasses and tabletop mirror, and engages with Child in trying on different items. Then Examiner puts pair of gloves on ears. If Child does not initiate JA, Examiner places hat on foot.
6	RJA	Examiner gets Child's attention (eye contact or visual orientation to Examiner). Then Examiner LOOKS and POINTS at a picture on wall to right.
7	IJA	*Wristband:* Examiner offers Child choice of puzzles. Takes pieces out of Child's chosen puzzle and gives pieces to Child 1 by 1 as child puts puzzle together. Rater first 2 pieces, Examiner discreetly slips on colorful, silly-looking wristband; offers child 3rd puzzle piece using hand with wristband. If Child doesn't initiate IJA, Examiner removes wristband, gives child 2 more puzzle pieces, and then puts wristband on again prior to giving child the 6th puzzle piece.
8	RJA	Examiner gets Child's attention (eye contact or visual orientation to Examiner). Then Examiner LOOKS and POINTS at a picture on wall to left.
9	IJA	*Book with altered pages:* Examiner brings out several children's books and offers them to Child. Child chooses one book to look at together with Examiner. Each book has four pages altered (scribbles on page, upside- down page, torn page and blank page.
10	RJA	Examiner gets Child's attention (eye contact or visual orientation to Examiner). Then Examiner LOOKS and POINTS at a picture on wall to right and VERBALIZES, "Look."
11	IJA	*Musical instruments:* Examiner offers Child a container of musical instruments. Examiner & Child remove instruments from box & try them out. Examiner remains silent.
12	RJA	Examiner gets Child's attention (eye contact or visual orientation to Examiner). Then Examiner LOOKS and POINTS at a picture on wall to left and VERBALIZES, "Look".
13	IJA	*Remote control car:* Examiner hands Child distractor toy. Remote control car is hidden under chair. Examiner activates car so that it rolls out into child's view. If Child does not initiate JA, Examiner waits 10 seconds, then activates car again.
14	RJA	Examiner gets Child's attention (eye contact or visual orientation to Examiner). Then Examiner LOOKS and POINTS at a picture on wall to right while VERBALIZING, "Look, (name object in picture)."
15	IJA	*Bag of toys:* Examiner presents Child with an opaque bag of six varied toys, saying, "Here. You can look in this bag." Examiner pretends to be busy doing something else, giving Child time to explore toys in the bag.
16	RJA	Examiner gets Child's attention (eye contact or visual orientation to Examiner). Then Examiner LOOKS and POINTS at a picture on wall to the left, while VERBALIZING, "Look, (name of object in picture)."

Notes: IJA = Initiate Joint Attention; RJA = Response to Joint Attention; JA = Joint Attention.

Table 2

Quality Standards for Psychometric Measure Evaluation

Quality Standard	Description
1. Describe the measure	• Settings and population for which the measure was designed (Stone, Coonrod, & Ousley, 2000). • Goal and purpose of the measure (Friberg, 2010)
2. Describe the examiner qualifications	• Characteristics of the examiners and raters, including level of experience (Kottner et al., 2011). • Assessor training required to administer the measure reliably. • Disclose if raters were blind to the child's diagnosis (Stone et al., 2000).
3. Explain test procedures	• Adequate procedural description for clinician to replicate testing environment (Friberg, 2010).
4. Describe the sample population	• Sample population used to obtain psychometric data and determine appropriateness of the test for particular children; include: geographic region, age, gender, ethnicity, socioeconomic status, and diagnostic/disability categories. • Recruitment method and justification of sample size. • A sample size of at least 100 children per relevant population group (e.g. gender, race, etc 2010).
5. Examine and interpret psychometric properties of reliability and validity	Reliability *Internal consistency* • Guidelines for interpreting Cronbach's coefficient alpha: ≤.70 = unacceptable, .70–.79 = fair, .80–.89 = good, ≥ .90 = excellent (Cicchetti and Sparrow, 1990). • Type and purpose of the measure considered when interpreting reliability (Bracken, 1987) *Test-retest and Interrater reliability* • Guidelines for interpreting statistical significance for both test-retest and interrater reliability, typically reported using Cohen's Kappa (Cohen, 1960; for nominal data) or an Intraclass Correlation Coefficient (ICC; for ordinal, interval, or ratio data): <.40 = poor, .40–.59 = fair, .60–.74 = good, .75–1.00 = excellent (Cicchetti and Sparrow, 1981). • A test-retest interval of 2–4 weeks recommended for young children (Bracken, 1987; Cicchetti, 1994). Validity *Concurrent and Divergent validity* • Correlation of 1.00 or −1.00 not expected or desired; new test should not be an exact replica of the other measure (Cicchetti, 1994; Crocker & Algina, 1986).
6. Examine the dimensionality of the measure	• Factor structure of the measure considered when evaluating subscales. • Construct validity supported with item total correlations greater than .30, inter-item correlations between .30 and .70, and factor analysis with factor loadings greater than .40 (DeVon et al., 2007).

Table 3

Sample 1 Demographic Information

Diagnosis	ASD (n= 122)	DD (n= 57)	TD (n= 81)
Chronological Age in Months			
M(*SD*)	61.18 (*25.58*)	55.82 (*40.42*)	50.90 (*29.04*)
Range	14–154	20–134	10–122
Gender			
% Male	83%	54%	75%
Race			
% Caucasian	80%	81%	80%
% African American	9%	7%	7%
% Other/Mixed Race	9%	11%	12%
% Undisclosed	2%	2%	1%
Mother's Level of Education Completed			
% High School or GED	23%	20%	5%
% College or Technical	56%	47%	47%
% Graduate School	21%	33%	48%
Household Income			
% <$60,000	43%	36%	39%
% ≥$60,000	56%	63%	61%

Table 4

Sample 1 Performance on Behavioral and Parent Report Measures Table

	ASD Sample			DD Sample			TD Sample		
Measure	**N**	**Mean (*SD*)**	**Range**	**N**	**Mean (*SD*)**	**Range**	**N**	**Mean (*SD*)**	**Range**
MSEL – Receptive Language	64	28.20 (11.88)	20–63	34	28.97 (10.04)	20–54	38	57 (7.79)	45–77
MSEL – Expressive Language	66	27.45 (10.63)	20–63	34	27.65 (8.62)	20–53	38	57.76 (9.21)	34–76
MSEL [*] – Early Learning Composite	63	59.27 (*15.81*)	49–99	41	60.90 (*13.32*)	49–103	47	112.34 (*11.80*)	84–147
Preschool Language Scales- Auditory Comprehension	67	70.15 (23.21)	50–126	36	72.89 (16.13)	50–116	50	113.12 (12.67)	81–145
Preschool Language Scales – Expressive Communication	67	68.27 (20.40)	50–119	36	74.06 (15.08)	50–122	50	116.02 (13.23)	79–145
Preschool Language Scales – Total Language Score	67	67.96 (*21.82*)	50–117	36	71.11 (*16.18*)	50–122	50	116.3 (*13.60*)	90–149
VABS [*]- Adaptive Behavior Composite	119	59.87 (*15.30*)	29–103	55	61 (*15.44*)	23–104	79	99.52 (*11.72*)	68–129

[*] MSEL = Mullen Scales of Early Learning (Mullen, 1995);

[*] VABS=Vineland Adaptive Behavior Scales (Sparrow et al., 1984)

Table 5

Sample 2 Demographic Information

Diagnosis	ASD (n=19)
Chronological Age in Months	
M(*SD*)	71.21 (*17.85*)
Range	37–94
Gender	100% Male
Race	58% White
	11% Black or African American
	5% Asian
	5% American Indian
	21% Other or Not Specified
SCQ* Total Score	
M(*SD*)	16.21 (*4.98*)
Range	9–28

*SCQ = Social Communication Questionnaire (Rutter et al., 2003)

Table 6

Test-retest ICCs for Sample 2

JA Protocol Item	ICC	Interpretation (Cicchetti & Sparrow, 1990)
RJA 1	.50	Fair
RJA 2	.62	Good
RJA 3	.65	Good
RJA 4	0*	Unknown
RJA 5	.65	Good
RJA 6	0*	Unknown
RJA 7	.65	Good
RJA 8	0*	Unknown
IJA 1	.55	Fair
IJA 2	.71	Good
IJA 3	.21	Poor
IJA 4	.36	Poor
IJA 5	.49	Fair
IJA 6	.62	Good
IJA 7	.18	Poor
IJA 8	.30	Poor
RJA Mean	.80	Good
IJA Mean	.83	Good

* 0 values due to errors in variance/covariance matrix when calculating ICCs for this small sample, likely due to perfect agreement on these items. Kappa statistics for these items = 1.

Table 7

Interrater Reliability for both Samples, Intraclass Correlation Coefficients

JA Protocol Item	Sample 1 ICC	Interpretation (Cicchetti and Sparrow, 1981)	Sample 2 ICC	Interpretation (Cicchetti and Sparrow, 1981)
RJA 1	.89	Excellent	.50	Fair
RJA 2	.73	Good	.62	Good
RJA 3	.78	Excellent	.65	Good
RJA 4	1	Excellent	1	Excellent
RJA 5	.94	Excellent	.65	Good
RJA 6	1	Excellent	0*	Unknown
RJA 7	.87	Excellent	.65	Good
RJA 8	.93	Excellent	0*	Unknown
IJA 1	.51	Fair	.51	Fair
IJA 2	.72	Good	.78	Excellent
IJA 3	.74	Good	.28	Poor
IJA 4	.70	Good	.39	Poor
IJA 5	.75	Excellent	.48	Poor
IJA 6	.75	Excellent	.62	Good
IJA 7	.77	Excellent	.20	Poor
IJA 8	.77	Excellent	.33	Poor
RJA Mean	.97	Excellent	.99	Excellent
IJA Mean	.89	Excellent	.98	Excellent

*0 values due to errors in variance/covariance matrix when calculating ICCs for this small sample, likely due to perfect agreement on these items. Kappa statistics for these items = 1.

Table 8

Sample 1 Spearman Non-Parametric Correlations between IJA and RJA Subscale Mean Scores and Other Measures by Diagnostic Category; Convergent and Divergent Validity

	ASD		DD		TD	
	IJA Mean	**RJA Mean**	**IJA Mean**	**RJA Mean**	**IJA Mean**	**RJA Mean**
Convergent Validity						
ADOS Pointing	−.51**	−.61**	−.26	−.41*		
ADOS Showing	−.46**	−.35**	−.35*	−.22		
ADOS Spontaneous IJA	−.39**	−.37**	−.05	−.01		
ADOS RJA	−.50**	−.80**	−.28	−.37**		
PLS 4 Auditory Comprehension	.76**	.85**	.43**	.54**	.26	.58**
PLS 4 Expressive Language	.74**	.81**	.32*	.56**	.26	.55**
Mullen Receptive Communication	.66**	.84**	.46**	.55**	.27	.56**
Mullen Expressive Communication	.60**	.80**	.36*	.56**	.23	.56**
Divergent Validity						
ADOS Unusual Sensory Interest	−.18	−.39**	−.15	−.04		
ADOS Hand/Finger Mannerisms	−.21*	−.34**	−.12	−.04		
ADOS Repetitive Behaviors	−.36**	−.29**	−.21	−.17		
ADOS RRB Calibrated Severity Score	−.28**	−.44**	−.24	−.16		

ADOS module items were coded from both Modules 1 and 2 on a 0–2 scale with 2 indicating more deficits in that skill area, ADOS data were not collected for TD children; PLS4 = Preschool Language Scales 4th edition, raw scores used in correlations; Mullen Scales of Early Learning, raw scores used in correlations; ADOS RRB = ADOS Restricted and Repetitive Behavior Calibrated Severity Score based on Hus et al., 2014.

*Significant at alpha level of .05;

**Significant at alpha level of .01

Table 9

Sample 1 Inter-item Tetrachoric Correlations plus Point-Biserial Correlations between Items and Subscale Scores.

	IJA1	IJA2	IJA3	IJA4	IJA5	IJA6	IJA7	IJA8	RJA1	RJA2	RJA3	RJA4	RJA5	RJA6	RJA7	RJA8	IJA Mean	RJA Mean
IJA1	1	.70	.75	.67	.70	.71	.61	.76	.54	.51	.65	.49	.56	.49	.57	.61	.80	.48
IJA2		1	.75	.54	.68	.67	.60	.78	.55	.59	.60	.65	.71	.64	.57	.69	.77	.52
IJA3			1	.63	.74	.66	.65	.75	.59	.58	.63	.60	.70	.55	.64	.68	.80	.52
IJA4				1	.61	.62	.47	.65	.35	.35	.45	.45	.46	.50	.42	.53	.76	.40
IJA5					1	.65	.66	.68	.40	.43	.52	.48	.53	.53	.50	.56	.82	.42
IJA6						1	.63	.79	.56	.48	.76	.55	.65	.63	.63	.52	.78	.49
IJA7							1	.71	.44	.48	.54	.41	.60	.55	.56	.50	.75	.43
IJA8								1	.39	.45	.60	.63	.56	.60	.57	.57	.84	.45
RJA1									1	.81	.78	.73	.78	.66	.77	.76	.49	.81
RJA2										1	.78	.77	.71	.71	.77	.73	.49	.79
RJA3											1	.83	.89	.84	.91	.75	.59	.78
RJA4												1	.90	.94	.93	.86	.52	.83
RJA5													1	.96	.92	.95	.59	.90
RJA6														1	.95	.90	.55	.83
RJA7															1	.86	.54	.87
RJA8																1	.57	.81
IJA Mean																	1	.56
RJA Mean																		1

Table 10

Sample 1 Factor Loadings from Confirmatory Factor Analysis

Item	Factor 1	Factor 2
RJA 1	1 (fixed)	0
RJA 2	1.026	0
RJA 3	1.052	0
RJA 4	.827	0
RJA 5	.937	0
RJA 6	1.016	0
RJA 7	.893	0
RJA 8	1.035	0
IJA 1	0	1 (fixed)
IJA 2	0	1.000
IJA 3	0	1.117
IJA 4	0	1.137
IJA 5	0	1.195
IJA 6	0	1.176
IJA 7	0	1.163
IJA 8	0	1.142