ASHA American Speech-Language-Hearing Association

AJSLP

Research Article

# Autistic Children's Irritability During Social Communication Assessments

Hannah L. Fipp-Rosenfield,[a] Rachel S. Levy,[a] Jeffrey M. Grauzer,[a] Aaron Kaat,[b] and Megan Y. Roberts[a]

[a] Roxelyn and Richard Pepper Department of Communication Sciences and Disorders, Northwestern University, Evanston, IL [b] Department of Medical Social Sciences, Feinberg School of Medicine, Northwestern University, Chicago, IL

ARTICLE INFO

Article History:
Received June 15, 2023
Revision received September 11, 2023
Accepted October 10, 2023

Editor-in-Chief: Erinn H. Finke
Editor: Amy L. Donaldson

https://doi.org/10.1044/2023_AJSLP-23-00216

ABSTRACT

**Purpose:** Accurate measurement of autistic children's social communication is critical for assessing skills, setting intervention goals, evaluating change over time, determining service eligibility, and determining classroom placement. There are various types of assessments, some of which use specific tasks to elicit social communication. Structured tasks may frustrate children, inadvertently elicit irritability, and have a cascading effect on their ability to communicate. To date, no studies have evaluated how differing types of social communication assessments may exacerbate children's irritability and impact assessment scores. We examined the extent to which (a) social communication assessment type (structured vs. naturalistic) impacts autistic children's irritability and (b) child irritability is associated with social communication scores.

**Method:** Autistic toddlers ($n$ = 114, $M_{age}$ = 33.09 months, $SD$ = 6.15) completed the Communication and Symbolic Behavior Scales (CSBS; structured) and a 10-min play-based mother–child interaction (MCX; naturalistic). Child irritability was scored on both assessments using a global rating scale of 0–15.

**Results:** Child irritability during the CSBS was significantly higher than during the MCX ($V$ = 4892, $p$ < .001, $r$ = .68). Higher irritability was associated with lower CSBS social communication scores ($B$ = −0.05, $p$ = .03), but not MCX scores ($B$ = 0.04, $p$ = .13; Theil's $F$ = 6.92, $p$ = .009).

**Conclusions:** Our findings suggest that the CSBS may pose unique challenges for autistic children, as it led to higher rates of irritability and negatively affected children's social communication scores. Evaluating the association between assessment type and irritability supports the complete characterization of autistic children's experience during assessments and clinicians in obtaining a more representative measure of social communication.

Accurate measurement of autistic children's social communication is a longstanding challenge (Izaryk et al., 2021; Wetherby, 2006). Within research and clinical settings, social communication assessments are used to measure autistic children's skills, set intervention goals, evaluate changes over time, and establish eligibility for services (Izaryk et al., 2021). There are various types of social communication assessments (e.g., structured, naturalistic), some of which use specific tasks to elicit children's social communication (Wetherby, 2006). Structured tasks may frustrate children, inadvertently elicit irritability, and/or have a cascading effect on their ability to communicate. Therefore, if an autistic child is irritable during a social communication assessment, it may hinder their ability to communicate and further impact their assessment score. The current study examines how different social communication assessments impact irritability in autistic children and whether irritability impacts their assessment scores. To our knowledge, this association has not been evaluated in any pediatric populations (i.e., autism, hearing loss,

Correspondence to Megan Y. Roberts: megan.y.roberts@northwestern.edu. ***Disclosure:*** *The authors have declared that no competing financial or nonfinancial interests existed at the time of publication.*

Down syndrome). As such, this analysis is a critical first step to determining factors that influence autistic children's social communication during various assessments.

Challenges with social communication are a core diagnostic feature of autism (American Psychiatric Association, 2013). Additionally, an estimated 50% of caregivers report that their autistic children display high rates of irritability (e.g., whining, throwing toys, tantrums), which has cascading effects on children's ability to communicate (Chow & Wehby, 2018; Kanne & Mazurek, 2011). Irritability is a common emotion seen across children's development that includes behavioral challenges such as verbal and physical aggression, self-injurious behaviors, and property destruction (Evans et al., 2017; Mikita et al., 2015; Robb, 2010). Often within toddlerhood, children will display irritability to communicate that they are upset, to express dysregulation, or as a result of frustration (Kaat et al., 2014). Previous studies have found a bidirectional association between autistic children's irritability and social communication (Chow & Wehby, 2018; Martínez-González et al., 2022) such that autistic children (a) display irritability when they are frustrated and unable to verbally communicate with a partner and (b) have less frequent social communication opportunities when they present higher levels of irritability (Boonen et al., 2014; Matson et al., 2009). Therefore, when evaluating autistic children's social communication, it is critical to conduct assessments when children are in an emotionally regulated and nonirritated state as the presence of irritability may hinder their ability to fully display their social communication skills (Martínez-González et al., 2022).

Despite the association between irritability and social communication, the effect of social communication assessments on autistic children's irritability has yet to be examined. There are many types of social communication assessments that vary based on social partner, interactive context, source of information, and psychometric features of measurement (Wetherby, 2006). Structured assessments are used to measure autistic children's social communication through standardized administration methods and systematic procedures typically consisting of communication probes or temptations and are scored based on children's norm-referenced age-matched peers (Wetherby, 2006; Wetherby & Prizant, 2003). These assessments may lead to increased child irritability, because some elicitation probes are administered by restricting access to test stimuli and waiting for the child to request or communicate. Responding to elicitation probes like these may be difficult for autistic children as social communication is a common challenge. Additionally, these assessments do not have a specified time length (e.g., administration time can vary from approximately 15 to 45 min), further exacerbating irritability due to testing fatigue. In contrast, naturalistic assessments evaluate children's social communication in a play-based context, typically with a caregiver, and are later scored based on frequency of the child's spontaneous communication productions (Wetherby, 2006). As these assessments have a fixed length of time, are play-based, and typically do not require the child to respond to probes or demands, they are less likely to lead to increased irritability.

Understanding the effects of social communication assessment on children's irritability is critical for speech-language pathologists' (SLPs) selection of assessment tools when evaluating young autistic children. SLPs use different social communication assessments for various purposes (e.g., eligibility for services, monitoring response to intervention). Within the literature, it is recommended that SLPs use structured assessments with caution as they lack ecological validity and are difficult to standardize (Izaryk et al., 2021; O'Neill, 2014; Timler, 2018). However, most states require clinicians to use at least one structured, norm-referenced measure when establishing a child's percent delay and determining eligibility criteria for special education services (Izaryk et al., 2021; Spaulding et al., 2012). Taken together, there appears to be a mismatch between social communication assessment recommendations for SLPs and the assessments that are required for eligibility services (Izaryk et al., 2021; Spaulding et al., 2012). Furthermore, if structured assessments lead to greater irritability, thereby affecting children's ability to demonstrate their social communication skills, assessment scores may not accurately represent the child's abilities.

Autistic stakeholders (e.g., parents of autistic children, autistic adults, the neurodiversity movement) encourage the evaluation of autistic children's experiences within research and clinical settings (Leadbitter et al., 2021). As such, investigating the association between social communication assessment type on child irritability and the cascading effect on assessment scores is one way to characterize the challenges autistic children may face when completing assessments (Solomon & Bagatell, 2010). Furthermore, valid measurement of autistic children's social communication is necessary for SLPs to accurately measure children's skills. The current study aims to evaluate the effects of (a) social communication assessment type (structured vs. naturalistic) on autistic children's irritability and (b) child irritability on social communication assessment scores. We hypothesized that the standardized probes and tasks of structured assessment may lead to frustration and the display of irritability, ultimately negatively impacting autistic children's assessment scores. In contrast, naturalistic assessment is less structured and play-based and may lead to comparatively lower irritability, thus less likely impacting autistic children's assessment scores.

## Method

### Study Design

The current study was a secondary analysis of data from a randomized clinical trial evaluating the effects of caregiver-implemented language facilitation strategies on the communication outcomes of autistic toddlers (R01DC014709, NCT02632773; Roberts et al., 2023). Families from the original clinical trial who had all required baseline data were included in the current study. Informed consent was obtained from all participants, and baseline data were collected before group randomization and 8-week intervention. Roberts et al. (2023) provided further information about intervention strategies. Participants were recruited continuously between July 2015 and March 2020.

### Participants

The current study was conducted in the greater Chicagoland area, and participants were recruited from pediatricians, autism diagnostic programs, and Part C Early Intervention providers. The Individuals with Disabilities Education Act (IDEA) Part C is a federal grant program that assists states in providing early intervention services for children with or at risk for disabilities between birth through 3 years of age (Individuals with Disabilities Education Act, 2004). Children were eligible for the study if they were between 18 and 48 months old; they had a diagnosis of autism or their mothers had concerns about autism, which were later confirmed by an in-person autism assessment including the Autism Diagnostic Observation Schedule–Second Edition (Lord et al., 2012); and they had no additional diagnosis that could impact language development. This study included 114 mother–child dyads (see Table 1). Children in the sample were predominately male ($n$ = 86, 75%) and had an average age of 33.09 months old ($SD$ = 6.15 months). The original clinical trial evaluated maternal characteristics (i.e., broader autism phenotype) on communication intervention strategy use. As such, caregivers were eligible for the clinical trial if they were the biological mother to the participating child; spoke English at least 50% of the time with the participating child; and did not have a diagnosis of fragile X syndrome, cerebral palsy, schizophrenia, profound hearing loss, or brain or head injury where she lost consciousness.

### Measures

#### Structured Measure of Social Communication

Before intervention at baseline, the Communication and Symbolic Behavior Scales (CSBS; Wetherby & Prizant, 2003) was administered by either a research assistant, a

**Table 1.** Participant demographics.

| | | |
|---|---|---|
| **Child characteristics** | | |
| Age | Months | 33.09 (6.15) |
| Biological sex | Male | 86 (75%) |
| | Female | 28 (25%) |
| Race | White | 59 (52%) |
| | Asian | 3 (3%) |
| | Black | 14 (12%) |
| | American Indian or Alaska Native | 1 (1%) |
| | More than one race | 26 (23%) |
| | Prefer not to answer | 9 (8%) |
| Ethnicity | Hispanic | 40 (35%) |
| | Not Hispanic or Latino | 68 (60%) |
| | Prefer not to answer | 4 (4%) |
| Child irritability | Coded on CSBS | 6.89 (4.47) |
| | Coded on MCX | 3.01 (3.55) |
| Naturalistic communication | Child's spontaneous and directed utterances on MCX | 6.42 (7.51) |
| Structured communication | CSBS | 57.52 (33.99) |
| **Mother characteristics** | | |
| Age | Years | 35.29 (5.18) |
| Race | White | 65 (57%) |
| | Asian | 7 (6%) |
| | Black | 16 (14%) |
| | Native Hawaiian or other Pacific Islander | 1 (1%) |
| | American Indian or Alaska Native | 2 (2%) |
| | More than one race | 8 (7%) |
| | Prefer not to answer | 11 (10%) |
| Ethnicity | Hispanic | 31 (27%) |
| | Not Hispanic or Latino | 76 (67%) |
| | Prefer not to answer | 3 (3%) |
| Education | Without HS | 2 (2%) |
| | HS graduate | 16 (14%) |
| | Some college | 28 (25%) |
| | College graduate from 4-year college or more | 65 (57%) |
| Employment status | Not employed | 4 (4%) |
| | Stay-at-home mother | 38 (33%) |
| | Part-time | 28 (25%) |
| | Full-time | 38 (33%) |
| | Second job | 3 (3%) |
| Income | < $30,000 | 11 (10%) |
| | $30,000–$49,000 | 13 (11%) |
| | $50,000–$100,000 | 24 (21%) |
| | > $100,000 | 46 (40%) |
| | No response | 17 (15%) |
| Government assistance | Receiving assistance | 23 (20%) |
| | Not receiving assistance | 88 (77%) |

*Note*. CSBS = Communication and Symbolic Behavior Scales; MCX = mother–child interaction; HS = high school.

PhD student, or an SLP. The CSBS is a structured measure that captures children's social communication. During the CSBS, children were presented with six sampling opportunities to elicit communication (i.e., wind-up toy, balloon, bubbles, snack/object in a container, books, and toys related to symbolic play) that yield seven cluster scores that correspond to social communication skills. Within each of the six sampling opportunities, children have various levels of access to materials. For example, within the snack/object in a container activity, the snack/object is sealed within a container and placed in front of the child. The examiner is instructed to look expectantly at the child and use two additional communication temptations (e.g., reach hand out and say, "Need help?") if the child does not respond. Finally, if the child did not respond and the examiner initiated the three standardized communication temptations, the child would gain access to the snack/object within the container. Although the CSBS is a structured measure, there are activities within the assessment that allow the child more access to materials while the clinician engages by following the child's lead (e.g., toy play with kitchen items within the object use section). A total raw score is derived from a sum of the seven cluster scores. Reliability was completed for 20% of CSBS administrations with .99 for intraclass correlations (ICCs; Koo & Li, 2016). Some children within our sample were outside the age of the normative sample for the CSBS (i.e., 24 months). As such, we used the CSBS total weighted raw score and controlled for age within our statistical models.

### Naturalistic Measure of Social Communication

Before intervention at baseline, a 10-min naturalistic mother–child interaction (MCX) using a standardized set of toys was video-recorded. Before filming, mothers were instructed to play with their child as they normally would at home. Based on the original clinical trial's measurement of social communication, research staff used an event-based social communication code from video recordings of the MCX to capture every child utterance that was spontaneous and directed (Roberts et al., 2023). Child utterances were coded as spontaneous if they were not prompted, imitated, or elicited by the adult. Additionally, child utterances were coded as directed if the child (a) made clear eye contact paired with a vocalization, gesture, or other symbolic form; (b) used a gesture that was directed to their caregiver (i.e., point, show, give); (c) referenced their caregiver by name or use of a pronoun; or (d) communicated as a direct result of the caregiver's prompt (Roberts et al., 2023). Coders were trained to research reliability by achieving 80% agreement or above with a master transcriber/coder on three consecutive MCX recordings (Yoder et al., 2018). Social communication on the MCX was then calculated based on the total sum of child's utterances that were spontaneous and directed. Reliability was completed for 20% of the MCXs (ICC = .98; Koo & Li, 2016).

### Child Irritability

All CSBS and MCX video recordings were also rated using a child irritability global rating scale that was adapted by The Language, Education, and Reading Neuroscience Lab at Northwestern University (LEARN; E. Norton, personal communication, October 22, 2021) from the Disruptive Behavior Diagnostic Observation Schedule (DB-DOS; Wakschlag, Briggs-Gowan, et al., 2008; Wakschlag, Hill, et al., 2008). The irritability rating scale consists of five independent items: intensity, predominance, ease of elicitation, rapid escalation, and difficult recovering (see Table 2). Each item is scored on a 4-point Likert-type scale (i.e., 0–3), with the sum of all five items representing an overall score. Total irritability scores ranged from 0 to 15, with "0" representing low levels of irritability.

Coders were PhD and master's students in the field of speech-language pathology and communication sciences and disorders. Coders had experience administering structured assessments (e.g., the Clinical Evaluation of Language Fundamentals–Fifth Edition; Wiig et al., 2013) and parent-mediated interventions (e.g., Project ImPACT and Research Units in Behavioral Intervention; Bearss et al., 2015; Ingersoll & Wainer, 2013; Wiig et al., 2013). None of the irritability coders administered the CSBS or video-recorded the MCX. They were trained to reliability with a standard set of consensus-rated videos. Coders were reliable once ratings on three consecutive videos met the following criteria compared to the training samples: (a) all five items were within 1 point, and (b) the overall score was within 3 points (Frost et al., 2020; Yoder et al., 2018). ICCs were used to evaluate the agreement between coders on the overall score (Hallgren, 2012). In addition, 20% of all MCX and CSBS video scores for child irritability were coded by a second master's student, and discrepancies were discussed during weekly coding meetings. Overall ICC for child irritability of MCX videos was .80, and for CSBS, it was .92 (Koo & Li, 2016).

## *Statistical Method*

To address our first research aim, a Wilcoxon signed-ranks test was conducted to test the effect of social communication assessment type on autistic children's irritability. Our child irritability data across both the CSBS and MCX were positively skewed. Therefore, a Wilcoxon signed-ranks test was appropriate as it is a nonparametric test that does not assume normally distributed data (Woolson, 2008).

Before evaluating our second aim, the effects of child irritability on social communication assessment score, we conducted a natural log transformation to both

**Table 2.** Child irritability macrocode.

| Macrocode items | Coding scale | Code definition |
|---|---|---|
| Irritability intensity | 0–3, with 0 representing low and 3 representing high levels of irritability | • Intensity is scored as the highest intensity level that the child reaches during the activity<br>• 0 = *No irritability*<br>• 1 = *Scowling, whining, or fussing*<br>• 2 = *Brief verbal or nonverbal irritability* (e.g., yelling, throwing a toy)<br>• 3 = *Loud, intense, and sustained irritability* |
| Irritability predominance | 0–3, with 0 representing low and 3 representing high levels of irritability | • Predominance is scored as the frequency of irritability during the activity<br>• 0 = *No irritability*<br>• 1 = *Irritability is present less than 25% of the time*<br>• 2 = *Irritability is present approximately 25%–50% of the time*<br>• 3 = *Irritability is present more than 50% of the time* |
| Ease of elicitation | 0–3, with 0 representing low and 3 representing high levels of irritability | • Ease of elicitation is scored as how the child responds to challenging (e.g., restricted access to a toy, directions, questions) and nonchallenging (e.g., naturalistic play) activities<br>• 0 = *No irritability*<br>• 1 = *Occasional irritability in response to challenging activities*<br>• 2 = *Irritability elicited multiple times in response to challenging activities*<br>• 3 = *Irritability elicited multiple times in response to nonchallenging activities OR irritability elicited multiple times in response to both nonchallenging and challenging activities* |
| Rapid escalation | 0–3, with 0 representing low and 3 representing high levels of irritability | • Rapid escalation is scored based on the rate at which the child's irritability escalates<br>• 0 = *No irritability*<br>• 1 = *Irritability builds gradually*<br>• 2 = *Irritability builds more rapidly OR the child displays both slow and rapid escalation of irritability*<br>• 3 = *Irritability rapidly escalates and comes "out of the blue" with no warning* |
| Difficulty recovering | 0–3, with 0 representing low and 3 representing high levels of irritability | • Difficulty recovering is scored as how quickly and easily the child recovers from an irritable state<br>• 0 = *No irritability*<br>• 1 = *Mild difficulty calming down without the support of an adult*<br>• 2 = *Moderate difficulty calming down with adult support*<br>• 3 = *Extreme difficulty calming down with adult support* |
| Total score | 0–15, with 0 representing low and 15 representing high levels of irritability. | • Sum of all five items (i.e., irritability intensity, irritability predominance, ease of elicitation, rapid escalation, and difficulty recovering) |

the CSBS total raw score and MCX spontaneous and directed utterances due to both variables having positively skewed distributions. After conducting log transformations, we centered and standardized our two dependent variables (i.e., CSBS total raw score and MCX social communication score), so they could be compared on the same scale. We then ran a correlation between our two dependent variables. We found that the CSBS total raw score and MCX social communication score had a statistically significant positive correlation ($r = .34$, $p < .001$). Due to the correlation between our two dependent variables, we built a seemingly unrelated regression (SUR) model to analyze the association between child irritability and social communication assessment scores while controlling for children's age (see Table 3). By utilizing a SUR model, we allow for two correlated dependent variables within one model; this approach is equivalent to having two separate linear models but is statistically more efficient (Zellner, 1962). After running the SUR model, we reviewed the correlation of the CSBS total raw score and

**Table 3.** Results from the seemingly unrelated regression (SUR) model evaluating the association between child irritability and social communication scores on the Communication and Symbolic Behavior Scales (CSBS) and the mother–child interaction (MCX).

| Independent variable | CSBS social communication score[a] | | MCX social communication score[b] | | Difference test[c] |
|---|---|---|---|---|---|
| | *B* | *SE* | *B* | *SE* | *F* |
| Intercept | –0.84 | 0.55 | –0.18 | 0.53 | |
| Age (months) | 0.03* | 0.01 | 0.002 | 0.02 | |
| Child irritability | –0.05* | 0.02 | 0.04 | 0.03 | 6.92** |

*Note.* *SE* = standard error.

[a]$R^2$ = .11. [b]$R^2$ = .02. [c]Linear hypothesis testing (Theil's *F* test) was conducted to determine whether the identified coefficients were statistically significantly different between the CSBS and MCX social communication models.

*$p < .05$. **$p < .01$.

MCX social communication score residuals ($r$ = .39). This small-to-moderate correlation suggests that there were efficiency gains by estimating the parameters of the two equations jointly using a SUR model as compared to two separate ordinary least squares models. Linear hypothesis testing using Theil's $F$ test was completed to determine whether identified coefficients were significantly different between the CSBS and MCX components within the SUR model. Finally, visual analysis revealed that the natural log transformations to our dependent variables met the statistical assumptions for SUR models, in particular normality of residuals and homoscedasticity.

## Results

The current study evaluated the effect of social communication assessment type (CSBS, structured vs. MCX, naturalistic) on autistic children's irritability. The results of a Wilcoxon signed-ranks test found that child irritability during the CSBS was significantly higher than during the MCX ($V$ = 4,892, $p < .001$, $r$ = .68). The mean irritability score was 6.89 on the CSBS and 3.01 on the MCX. These findings support our hypothesis that communication elicitation probes within the CSBS are associated with higher levels of irritability as compared to naturalistic play during the MCX.

A SUR model was conducted to evaluate the effect of child irritability on social communication assessment score (see Table 3). Within the CSBS component of the SUR model, we found that higher child irritability was associated with lower CSBS total raw scores ($B$ = −0.05, $p$ = .03). Within the MCX component of the SUR model, there was no significant effect of child irritability on the MCX social communication scores ($B$ = 0.04, $p$ = .13). Results of Theil's $F$ linear hypothesis test allowed us to compare effects from the CSBS and MCX components and found that child irritability had a significantly larger effect on social communication scores of the CSBS than those of the MCX ($F$ = 6.92, $p$ = .009). As such, we found that higher levels of child irritability affect children's social communication abilities, subsequently impacting their social communication assessment scores, particularly in more structured assessment contexts such as the CSBS. Specifically, we see from Theil's $F$ test that there is a larger effect on social communication scores of the CSBS than the MCX, suggesting that irritability on the CSBS is more likely to impact children's social communication abilities and assessment scores as compared to the MCX.

## Discussion

The goal of the current study was to examine the effects of social communication assessment type on autistic children's irritability and how irritability affects social communication assessment score. Our findings indicate that the structured CSBS led to higher irritability in autistic children as compared to the naturalistic MCX. Additionally, child irritability on the CSBS was associated with lower social communication scores. Taken together, our findings provide preliminary evidence that more structured assessments like the CSBS may not fully capture autistic children's social communications skills.

We found that irritability on the CSBS led to lower assessment scores, which may have cascading effects on children's qualification for intervention services, intervention goals, and classroom placement. Most states require at least one structured, norm-referenced assessment score to establish a child's percent delay, compare their score to age-matched peers, determine early intervention eligibility, and later determine classroom placement (Izaryk et al., 2021; O'Neill, 2014; Timler, 2018). Under IDEA (2004), it is mandated that students with disabilities are evaluated and placed within the least restrictive environment (LRE). However, if autistic students are evaluated using structured, norm-referenced assessments, it may lead to higher

irritability and lower social communication assessment scores. As such, inaccurate social communication assessment scores may lead to overidentification of social communication difference, misinform classroom placement, and incorrect assigned classrooms and levels of support (e.g., being placed within a special education classroom instead of an LRE such as a general education classroom with a support aide). Furthermore, structured assessments are typically only administered once or twice a year because of potential measurement bias (e.g., practice effects) due to children's ability to recall stimuli and administration methods (Izaryk et al., 2021; Wetherby & Prizant, 2003). Therefore, eligibility services and classroom placement may be heavily dependent on only one standardized observation of autistic children's social communication skills. Additional research is needed to assess the ecological validity of structured social communication assessments, as they may not accurately represent autistic children's social communication skills and may negatively impact children's intervention eligibility, services, and later classroom placement (Izaryk et al., 2021).

Although we found that structured assessments led to higher rates of autistic children's irritability, future research is needed to evaluate what additional variables may be driving the association between assessment type and irritability. For example, familiarity with the adult administering the assessment may impact how irritable children become. Within the current study, the CSBS assessor was an adult unfamiliar to the child, while the MCX was completed with the child's mother. It may be that the child's mother is more familiar with their child's temperament. Therefore, the mother may be able to respond to the child's irritability and facilitate emotional regulation in times of distress as compared to an unfamiliar assessor. Additionally, within the CSBS, it may be that the structured probes are eliciting higher rates of irritability as compared to the less structured play (e.g., toy play with kitchen items within the object use section). Future research could use an observational microcode for irritability to conduct a sequential analysis and determine the behavior or activity that immediately precedes child irritability when completing structured and naturalistic assessments.

By evaluating the associations between assessment type and child irritability, our findings suggest several key clinical takeaways when measuring social communication and considering autistic children's lived experiences during assessments. Autistic stakeholders have asked researchers and clinicians to consider how clinical activities impact children's mood, emotional regulation, and behavior (Leadbitter et al., 2021). Some structured assessments may frustrate autistic children, as they conclude after a fixed number of incorrect responses as compared to a set time (e.g., 10-min naturalistic observation). As such, longer assessments may lead to testing fatigue and result in frustration and irritability, thus negatively impacting autistic children's experience. However, clinicians will likely need to administer structured, norm-referenced tools as many are required for children's eligibility for special education services (Izaryk et al., 2021). SLPs may supplement structured, norm-referenced assessments with additional assessments (e.g., parent report, naturalistic assessments, language, communication samples) that are less frustrating for children and help fully understand children's social communication profile and inform their intervention services.

One potential benefit of administering both structured and naturalistic assessments is that it captures a more holistic picture of the child's skills. By measuring skills across various contexts, SLPs can observe when the child displays their highest communication abilities and irritability during challenging activities (e.g., structured play, communication probes and temptations). By evaluating the child's range of skills, SLPs may be able to identify the probes that elicit irritability to help target specific areas of social communication. For example, if the child cries when presented with a snack in a sealed jar, an SLP may consider targeting the child's use of functional communication to request wants and express needs. Additional research should consider evaluating how SLPs can use multiple assessments to establish individualized communication therapy goals for autistic children. Limitations to the current study include only analyzing the effects of one type of structured and naturalistic assessment on autistic children's irritability, which restricts our conclusions to only the effects of the CSBS and a naturalistic mother–child play assessment. Future studies should evaluate other types of naturalistic and structured assessments to determine their effects on child irritability. Furthermore, additional research should consider how different types of assessments beyond just structured and naturalistic (e.g., norm referenced, criterion referenced, language samples) affect child irritability. An additional limitation is the collection of only one observation for each assessment type as child irritability is context dependent and may vary based on the child's emotional state (Yoder et al., 2018). Therefore, future studies should consider how confounding variables such as child mood or fatigue may affect their irritability when completing assessments. Additionally, our study did not evaluate the effects of child's social communication skills on irritability over time. Finally, our sample has a larger proportion of boys as compared to girls. Future research may consider oversampling for girls as some research has identified differences in irritability based on biological sex (Humphreys et al., 2019; Leibenluft et al., 2006). Despite these limitations, our findings suggest the CSBS may not completely capture autistics toddlers' social communication skills, due to

significantly higher rates of irritability as compared to a naturalistic assessment.

The results of the present study extend prior research by examining the effects of social communication assessment type on autistic children's irritability and social communication score. We found that the structured CSBS led to higher rates of autistic children's irritability as compared to the naturalistic MCX. Additionally, child irritability on the CSBS was associated with lower assessment scores. Structured, norm-referenced assessments are often needed for early intervention eligibility services and later classroom placement; however, they may not accurately depict autistic children's social communication skills. As such, clinicians should consider supplementing structured assessments with additional measures (i.e., naturalistic observations) to fully understand autistic children's social communication skills.

## Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon request.

## Acknowledgments

This study was a secondary analysis from a randomized clinical trial funded by the National Institutes of Health, investigating the effects of caregiver-implemented communication facilitation strategies on autistic children's communication outcomes (R01DC014709, NCT02632773). The authors would like to thank the Early Intervention Research Group clinical staff and study participants. Without their participation and support, this project would not be possible.

## References

**American Psychiatric Association.** (2013). *Diagnostic and statistical manual of mental disorders: DSM-5* (5th ed.). https://doi.org/10.1176/appi.books.9780890425596

**Bearss, K., Johnson, C., Smith, T., Lecavalier, L., Swiezy, N., Aman, M., McAdam, D. B., Butter, E., Stillitano, C., Minshawi, N., Sukhodolsky, D. G., Mruzek, D. W., Turner, K., Neal, T., Hallett, V., Mulick, J. A., Green, B., Handen, B., Deng, Y., & Scahill, L.** (2015). Effect of parent training vs parent education on behavioral problems in children with autism spectrum disorder: A randomized clinical trial. *JAMA, 313*(15), 1524–1533. https://doi.org/10.1001/jama.2015.3150

**Boonen, H., Maljaars, J., Lambrechts, G., Zink, I., Van Leeuwen, K., & Noens, I.** (2014). Behavior problems among school-aged children with autism spectrum disorder: Associations with children's communication difficulties and parenting behaviors. *Research in Autism Spectrum Disorders, 8*(6), 716–725. https://doi.org/10.1016/j.rasd.2014.03.008

**Chow, J. C., & Wehby, J. H.** (2018). Associations between language and problem behavior: A systematic review and correlational meta-analysis. *Educational Psychology Review, 30*(1), 61–82. https://doi.org/10.1007/s10648-016-9385-z

**Evans, S. C., Burke, J. D., Roberts, M. C., Fite, P. J., Lochman, J. E., de la Peña, F. R., & Reed, G. M.** (2017). Irritability in child and adolescent psychopathology: An integrative review for ICD-11. *Clinical Psychology Review, 53,* 29–45. https://doi.org/10.1016/j.cpr.2017.01.004

**Frost, K. M., Brian, J., Gengoux, G. W., Hardan, A., Rieth, S. R., Stahmer, A., & Ingersoll, B.** (2020). Identifying and measuring the common elements of naturalistic developmental behavioral interventions for autism spectrum disorder: Development of the NDBI-Fi. *Autism, 24*(8), 2285–2297. https://doi.org/10.1177/1362361320944011

**Hallgren, K. A.** (2012). Computing inter-rater reliability for observational data: An overview and tutorial. *Tutorials in Quantitative Methods for Psychology, 8*(1), 23–34. https://doi.org/10.20982/tqmp.08.1.p023

**Humphreys, K. L., Schouboe, S. N. F., Kircanski, K., Leibenluft, E., Stringaris, A., & Gotlib, I. H.** (2019). Irritability, externalizing, and internalizing psychopathology in adolescence: Cross-sectional and longitudinal associations and moderation by sex. *Journal of Clinical Child & Adolescent Psychology, 48*(5), 781–789. https://doi.org/10.1080/15374416.2018.1460847

**Individuals with Disabilities Education Act of 2004, 20 U.S.C. § 1400.** (2004). https://sites.ed.gov/idea/

**Ingersoll, B., & Wainer, A.** (2013). Initial efficacy of Project ImPACT: A parent-mediated social communication intervention for young children with ASD. *Journal of Autism and Developmental Disorders, 43*(12), 2943–2952. https://doi.org/10.1007/s10803-013-1840-9

**Izaryk, K., Edge, R., & Lechwar, D.** (2021). A survey of speech-language pathologists' approaches to assessing social communication disorders in children. *Perspectives of the ASHA Special Interest Groups, 6*(1), 1–17. https://doi.org/10.1044/2020_PERSP-20-00147

**Kaat, A. J., Lecavalier, L., & Aman, M. G.** (2014). Validity of the aberrant behavior checklist in children with autism spectrum disorder. *Journal of Autism and Developmental Disorders, 44*(5), 1103–1116. https://doi.org/10.1007/s10803-013-1970-0

**Kanne, S. M., & Mazurek, M. O.** (2011). Aggression in children and adolescents with ASD: Prevalence and risk factors. *Journal of Autism and Developmental Disorders, 41*(7), 926–937. https://doi.org/10.1007/s10803-010-1118-4

**Koo, T. K., & Li, M. Y.** (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine, 15*(2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012

**Leadbitter, K., Buckle, K. L., Ellis, C., & Dekker, M.** (2021). Autistic self-advocacy and the neurodiversity movement: Implications for autism early intervention research and practice. *Frontiers in Psychology, 12,* Article 635690. https://doi.org/10.3389/fpsyg.2021.635690

**Leibenluft, E., Cohen, P., Gorrindo, T., Brook, J. S., & Pine, D. S.** (2006). Chronic versus episodic irritability in youth: A community-based, longitudinal study of clinical and diagnostic associations. *Journal of Child and Adolescent Psychopharmacology, 16*(4), 456–466. https://doi.org/10.1089/cap.2006.16.456

**Lord, C., Rutter, M., DiLavore, P. C., Risi, S., Gotham, K., & Bishop, S.** (2012). *Autism Diagnostic Observation Schedule–Second Edition (ADOS-2)*. Western Psychological Services.

**Martínez-González, A. E., Cervin, M., & Piqueras, J. A.** (2022). Relationships between emotion regulation, social communication and repetitive behaviors in autism spectrum disorder. *Journal of Autism and Developmental Disorders, 52*(10), 4519–4527. https://doi.org/10.1007/s10803-021-05340-x

**Matson, J. L., Boisjoli, J., & Mahan, S.** (2009). The relation of communication and challenging behaviors in infants and toddlers with autism spectrum disorders. *Journal of Developmental and Physical Disabilities, 21*(4), 253–261. https://doi.org/10.1007/s10882-009-9140-1

**Mikita, N., Hollocks, M. J., Papadopoulos, A. S., Aslani, A., Harrison, S., Leibenluft, E., Simonoff, E., & Stringaris, A.** (2015). Irritability in boys with autism spectrum disorders: An investigation of physiological reactivity. *Journal of Child Psychology and Psychiatry, 56*(10), 1118–1126. https://doi.org/10.1111/jcpp.12382

**O'Neill, D. K.** (2014). Assessing pragmatic language functioning in young children: Its importance and challenges. In D. Matthews (Ed.), *Pragmatic development in first language acquisition* (pp. 363–386). John Benjamins. https://doi.org/10.1075/tilar.10.20nei

**Robb, A. S.** (2010). Managing irritability and aggression in autism spectrum disorders in children and adolescents. *Developmental Disabilities Research Reviews, 16*(3), 258–264. https://doi.org/10.1002/ddrr.118

**Roberts, M. Y., Sone, B. J., Jones, M., Grauzer, J., Sudec, L., Stern, Y. S., Kwok, E., Losh, M., & Kaat, A.** (2023). One size does not fit all for parent-mediated autism interventions: A randomized clinical trial. *Autism, 27*(2), 443–455. https://doi.org/10.1177/13623613221102736

**Solomon, O., & Bagatell, N.** (2010). Introduction: Autism: Rethinking the possibilities. *Ethos, 38*(1), 1–7. https://doi.org/10.1111/j.1548-1352.2009.01078.x

**Spaulding, T. J., Swartwout Szulga, M., & Figueroa, C.** (2012). Using norm-referenced tests to determine severity of language impairment in children: Disconnect between U.S. policy makers and test developers. *Language, Speech, and Hearing Services in Schools, 43*(2), 176–190. https://doi.org/10.1044/0161-1461(2011/10-0103)

**Timler, G. R.** (2018). Similar … but very different: Determining when a child has social communication disorder versus autism spectrum disorder can be tricky. Here are some key considerations. *The ASHA Leader, 23*(4), 56–61. https://doi.org/10.1044/leader.FTR2.23042018.56

**Wakschlag, L. S., Briggs-Gowan, M. J., Hill, C., Danis, B., Leventhal, B. L., Keenan, K., Egger, H. L., Cicchetti, D., Burns, J., & Carter, A. S.** (2008). Observational assessment of preschool disruptive behavior, Part II: Validity of the Disruptive Behavior Diagnostic Observation Schedule (DB-DOS). *Journal of the American Academy of Child & Adolescent Psychiatry, 47*(6), 632–641. https://doi.org/10.1097/CHI.0b013e31816c5c10

**Wakschlag, L. S., Hill, C., Carter, A. S., Danis, B., Egger, H. L., Keenan, K., Leventhal, B. L., Cicchetti, D., Maskowitz, K., Burns, J., & Briggs-Gowan, M. J.** (2008). Observational assessment of preschool disruptive behavior, Part I: Reliability of the Disruptive Behavior Diagnostic Observation Schedule (DB-DOS). *Journal of the American Academy of Child & Adolescent Psychiatry, 47*(6), 622–631. https://doi.org/10.1097/CHI.0b013e31816c5bdb

**Wetherby, A. M.** (2006). Understanding and measuring social communication in children with autism spectrum disorders. In T. Charman & W. Stone (Eds.), *Social and communication development in autism spectrum disorders: Early identification, diagnosis, and intervention* (pp. 3–34). Guilford Press.

**Wetherby, A. M., & Prizant, B. M.** (2003). *Communication and Symbolic Behavior Scales (CSBS), Normed Edition*. Brookes.

**Wiig, E. H., Semel, E., & Secord, W. A.** (2013). Clinical evaluation of language fundamentals–fifth edition (CELF-5). *Journal of Psychoeducational Assessment, 33*(5), 495–500.

**Woolson, R. F.** (2008). Wilcoxon signed-rank test. In R. B. D'Agostino, L. Sullivan, & J. Massaro (Eds.), *Wiley encyclopedia of clinical trials*. Wiley. https://doi.org/10.1002/9780471462422.eoct979

**Yoder, P. J., Lloyd, B., & Symons, F. J.** (2018). *Observational measurement of behavior* (2nd ed.). Brookes.

**Zellner, A.** (1962). An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias. *Journal of the American Statistical Association, 57*(298), 348–368. https://doi.org/10.1080/01621459.1962.10480664