

# Conducting oral and written language adapted tele-assessments with early elementary-age children with autism spectrum disorder

Autism & Developmental Language Impairments  
Volume 7: 1–15  
© The Author(s) 2022  
Article reuse guidelines:  
sagepub.com/journals-permissions  
DOI: 10.1177/23969415221133268  
journals.sagepub.com/home/dli



Carlin Conner , Alyssa R Henry<sup>#</sup> and Emily J Solari

University of Virginia, Charlottesville, VA, USA

Matthew C Zajic

Teachers College, Columbia University, New York, NY, USA

## Abstract

**Background and aims:** Due to the COVID-19 pandemic, tele-health has gained popularity for both providing services and delivering assessments to children with disabilities. In this manuscript, we discuss the process of collecting standardized oral language, reading, and writing tele-assessment data with early elementary children with autism spectrum disorder (ASD) and offer preliminary findings related to child and parent engagement and technology issues.

**Methods:** The data presented are from pretest assessments during an efficacy study examining the electronic delivery of a listening comprehension intervention for children with ASD. Pretest sessions included a battery of standardized language, reading, and writing assessments, conducted over Zoom. The authors operationalized and developed a behavioral codebook of three overarching behavioral categories (parent involvement, child disengagement, and technology issues). Researchers coded videos offline to record frequencies of indicated behaviors across participants and assessment subtests.

**Results:** Involvement from parents accounted for the highest number of codes. Children showed some disengagement during assessment sessions. Technology issues were minimal. Behavioral categories appeared overall limited but varied across participants and assessments.

**Conclusions:** Parent involvement behaviors made up approximately two-thirds of the coded behaviors. Child disengagement behaviors made up approximately one-fourth of the coded behaviors, and these behaviors occurred more frequently across many different participants (with lower frequencies but greater coverage across children). Technology problems specific to responding to assessment items were relatively uncommon.

**Implications:** Clear guidelines including assessment preparation, modification of directions, and guidelines for parents who remain present are among the implications discussed. We also provide practical implications for continued successful adapted tele-assessments for children with ASD.

## Keywords

Autism spectrum disorder, tele-assessment, standardized assessment, oral language, reading, writing

In recent years, tele-health has gained popularity as a reliable tool through which services and assessments for children with disabilities can be delivered (Sutherland et al.,

2017, 2019; Waite et al., 2010). However, the popularity of tele-health services grew dramatically in March of 2020 due to the COVID-19 pandemic, as many professions

<sup>#</sup>Dr. Henry is currently at Texas A&M University, College Station, TX, USA.

### Corresponding author:

Carlin Conner, Department of Education and Human Development, University of Virginia, Charlottesville, VA, USA.

Email: clc6wc@virginia.edu



across education, health, and human services were forced to make a quick transition to tele-health services to meet the needs of their students and clients. Platforms for delivering tele-health services include face-to-face live video chatting, many with share-screen features enabling the practitioner to present information from their screen to the client, and chat box features for typing answers, if necessary. School professionals and educational researchers were also required to adapt assessments designed for in-person implementation for delivery over technology, something for which practitioners are continuing to seek guidance (Krach et al., 2020; Pearson Education, 2020).

From the outset of the pandemic, limited guidance existed for adapting assessments to be appropriately delivered over virtual platforms (Wright, 2020); furthermore, available guidance appeared conflicting (Krach et al., 2020). The translation of assessment materials and protocols designed for face-to-face administration to be used for tele-assessment has been termed adapted tele-assessment, and the use of adapted tele-assessment raises the concern for testing errors due to using such assessments in novel contexts that differ from how they were standardized and developed (Krach et al., 2020; Pearson Education, 2020).

In a national survey of families of children with autism spectrum disorder (ASD) conducted at the start of the COVID-19 pandemic ( $n = 3,502$ ), parents and caregivers indicated that by April 2020, many special education (47%) or speech and language therapy (50%) services for their children had been transferred to an online or remote format (White, Law, et al., 2021). As the pandemic continued, many families of children with ASD made use of tele-health services for the first time (White, Stoppelbein, et al., 2021). However, a limited amount of this research on the effectiveness of tele-assessment has included children with ASD or other neurodevelopmental disabilities (Sutherland et al., 2019). Research conducted during the COVID-19 pandemic with families of children with ASD reported positive feelings towards using tele-health services, yet cite concerns about effectiveness due to children's social communication difficulties (White, Stoppelbein, et al., 2021). Existing literature on the use of norm-referenced, psychoeducational assessments via tele-administration is limited (concerning either adapting existing assessments or developing new assessments to be used remotely during the COVID-19 pandemic). Psychoeducational language assessments have received little systematic attention, particularly with populations of children at-risk for language difficulties, including children with ASD, who often have sensory and/or social communication difficulties that may impact response to assessment (Ketterlin-Geller, 2008). Individuals with ASD are also more affected by disruptions or changes in schedule than their neurotypical peers, which may also affect assessment responses (White et al., 2021).

Systematic inquiry is needed to understand the nuanced challenges of adapting tests for different contexts and populations of learners (Pearson Education, 2020). This is especially important for assessments that are related to schooling and special education. Many children with ASD struggle with reading, particularly reading comprehension (McIntyre et al., 2017; Nation et al., 2006) and writing (Zajic & Wilson, 2020); therefore, accurate and reliable tele-assessment of these constructs will be useful to their assessors and service providers. Prior studies on the delivery and assessment of reading and writing instruction over technology platforms found either advantages or no difference in response to technology delivered services when compared to in-person services. Sutherland et al. (2019) found no significant differences between scores of norm-referenced language assessments in face-to-face compared to tele-health conditions for a small group of children with ASD ( $n = 13$ ; ages = 9–12). Pennington et al. (2020) noted no group-level differences in behavior between conditions but noted large variation between individuals, also common with face-to-face assessment of children with ASD. Regina Molini-Avejonas et al. (2015) systematically reviewed 103 articles focused on the delivery of speech and hearing services over a technology platform, finding more studies focused on hearing (32.1%) and speech (19.4%) than on language (16.5%). Included studies focused equally on assessment and intervention, with many finding advantages to delivering services via a technology platform. However, only three studies included participants with ASD and all outcome measures focused on parent feasibility and satisfaction (Baharav & Reiser, 2010; Vismara et al., 2012, 2013).

## Purpose

In this manuscript, we discuss the process of collecting standardized academic task data (oral language, reading, and writing assessments) via an adapted tele-assessment context (Zoom) with early elementary children with ASD during the COVID-19 pandemic. We were guided by the following research questions: (1) What is the nature of parent/caregiver involvement during adapted tele-assessment sessions of children with ASD?; (2) Do children with ASD disengage during adapted tele-assessment sessions, and if so, how often?; (3) Do technology issues interfere with assessment administration, and if so, how often?; and (4) Is the frequency of parent/caregiver involvement, child disengagement, and technology issues related to child characteristics (i.e., level of autism symptomatology and adaptive behavior skills)? In answering these research questions, we discuss the successes and challenges of conducting virtual assessments and provide practical implications to promote successful administration of adapted tele-assessments for elementary children with ASD.

**Table 1.** Sample descriptive statistics ( $n = 13$ ).

Variable	Scores		Administration duration (minutes)	
	Mean (SD)	Range	Mean (SD)	Range
Age (Years:Months)	7:11 (0:9)	6:8–9:8	–	–
Vineland-3 Adaptive Behavior Composite Standard Score	77.84 (7.75)	70–92	–	–
SRS-2 Parent Report T-Score	74.15 (7.51)	58–87	–	–
EVT-3 Standard Score	92.62 (16.77)	69–128	13.50 (5.41)	4.81–27.79
CELF-5 Word Structure Scaled Score	8.33 (3.61)	3–19	8.17 (2.78)	3.73–12.53
CELF-5 Understanding Spoken Paragraphs Scaled Score	5.85 (2.68)	2–9	7.82 (2.71)	3.73–13.38
NEPSY-II Narrative Memory Scaled Score	3.62 (1.90)	1–7	6.15 (1.77)	3.99–10.13
CTOPP-2 Rapid Digit Naming Scaled Score	6.85 (1.79)	3–10	1.82 (0.68)	1.23–3.85
WJ IV Letter-Word Identification Standard Score	91.38 (17.47)	55–128	3.01 (1.47)	1.19–5.41
WJ IV Word Attack Standard Score	91.23 (19.70)	40–123	3.64 (2.04)	1.84–8.00
WJ IV Sentence Writing Fluency Standard Score	76.08 (16.32)	47–104	10.73 (3.56)	3.45–16.41
WJ IV Writing Samples Standard Score	80.15 (12.25)	57–97	9.22 (6.18)	4.06–21.93

Note. Narrative Memory = Free and Cued Recall Scaled Score; SRS-2 = Social Responsiveness Scale, Second Edition; EVT-3 = Expressive Vocabulary Test, Third Edition; CELF-5 = Clinical Evaluation of Language Fundamentals, Fifth Edition; NEPSY-II = A Developmental NEuroPSYchological Assessment, Second Edition; CTOPP-2 = Comprehensive Test of Phonological Processing, Second Edition; WJ IV = Woodcock-Johnson Tests of Achievement, Fourth Edition; SD = standard deviation.

## Method

The data presented in this paper were part of an intervention feasibility study examining the virtual delivery of a listening comprehension intervention (building vocabulary and early reading strategies; Henry & Solari, 2020; Solari et al., 2020) for children with ASD. All data presented in this paper were collected prior to intervention during the pretest assessment battery.

### *Institutional review board approval, recruitment, and consent*

Following approval from the university Institutional Review Board, participants were recruited using a university-based ASD research initiative's system for individuals, families, and professionals affected by ASD. A flier was advertised via the initiative's Facebook page and newsletters as well as other social media outlets (i.e., Twitter, LinkedIn). To meet eligibility criteria, children must (1) have been diagnosed with ASD, (2) have been currently in 1st to 3rd grade, and (3) have adaptive behavior scores at or above 70 according to the Vineland Adaptive Behavior Scales, Third Edition (Vineland-3; Sparrow et al., 2016). The Vineland-3 is a norm-referenced comprehensive measure used to assess adaptive functioning. The adaptive behavior requirement was set in place to ensure that children had the ability to participate and attend to the standardized assessments. Prior to data collection, parents completed the Vineland-3 Domain-Level Parent/Caregiver Form and a simple demographic form to determine whether their child might be excluded based on other criteria, including significant psychiatric, sensory, or

motor impairments or any major medical disorder that could be associated with extended absences from school. Parent consent and child assent were obtained from all participants prior to beginning data collection.

### *Participant characteristics*

The recruitment period for this study was one month. Twenty-three caregivers expressed interest in the study and asked for additional information. Seventeen children (74%) were consented and assessed, and 14 (82% of those assessed) were determined to be eligible. Three children with ASD were excluded from the study due to adaptive behavior skills lower than 70. Another child was excluded from the present study due to missing data on some of the reading subtests. A total of 13 children with ASD met criteria for the study and had full pretest data available. The mean age of participating children was 7:11 (Range = 6:8–9:8). Two children were in 1st grade at the start of the intervention, 7 were in 2nd grade, and 5 were in 3rd grade. All parents provided proof of an ASD diagnosis via a letter from a pediatrician or a licensed psychologist, or confirmation of school services received under the classification of Autism. See Table 1 for participant information and pretest scores. The sample of children included 11 males and 2 females. Though group-level Vineland-3 Adaptive Behavior Composite (ABC) scores were below average ( $M = 77.84$ ,  $SD = 7.75$ ), no children had ABCs that were more than two standard deviations below the mean (i.e., all ABCs  $> 70$ ). Autism symptomatology, as measured by the Social Responsiveness Scale, Second Edition (SRS-2; Constantino & Gruber, 2012) was generally in the moderate range ( $M = 74.15$ ,  $SD =$

7.51; Range = 58–87). The sample included 7 White, 2 Hispanic/Latinx, 1 Asian, 1 Black, and 2 bi- or multi-racial participants. Four participants had a home language other than or in addition to English; however, no students were classified as Language Learners or provided with English Second Language services at school, thus, assessments were only conducted in English. In addition to ASD, participating children had diagnoses of speech sound disorder ( $n = 7$ ), attention-deficit/hyperactivity disorder ( $n = 6$ ), anxiety ( $n = 1$ ), and a specific learning disability ( $n = 1$ ). Eight children in the sample spent the majority of the day (at least 80%) in the general education classroom.

### Operationalizing and behavioral coding test session behaviors

To characterize behaviors during adapted tele-assessment sessions, the authors operationalized a behavioral codebook focused on events that could occur during the adapted tele-assessment sessions. All test sessions were recorded via Zoom and coded offline using Behavioral Observation Research Interactive Software (BORIS), an open-source event logging software for video coding (Friard et al., 2016). The third author trained two research assistants (one undergraduate student and one master's student) through a series of weekly meetings to establish conceptual understanding of the assessment sessions and the codebook. Training occurred over a two-month period with weekly discussions and minor modifications to the codebook. All test session videos were coded by one research assistant with a second research assistant coding all test sessions from a randomly selected set of seven participants (54%). Following independent coding, interrater reliability was calculated using the Gwet agreement coefficient for first-order chance correction ( $AC_1$ ; Gwet, 2014) via the kappaetc module in StataSE 17 (Klein, 2018). Table 2 provides  $AC_1$  coefficients for behavioral categories and individual behaviors.

The coding team took a multistep, systematic approach to exhaustively code all test session videos. First, the coding team collaboratively conducted a structural coding pass that demarcated out all test session tasks, including labeling sections by function (e.g., session explanation to participants, presentation of assessment materials for selected assessment subtests, transition between assessment activities). For each assessment administration, the coding team indicated sections for practice and training items as well as administered test items using duration codes. These task structure codes were checked across coders before coding behaviors to ensure agreement among coders for task structure. Second, the coding team conducted multiple passes of the test session videos coding for behaviors related to the individuals present during the test session (i.e., the parent and/or other adult present

with the child, the child, and the assessor). All behaviors were recorded as frequencies embedded within the task structure duration codes to account for the occurrence of behaviors within different sections of the assessment administration as well as different assessments. In general, behaviors could only be recorded once per assessment item unless another behavior disrupted that behavior either from the same individual (i.e., a parent offering corrective feedback, then offering a reinforcer, and then offering corrective feedback again within the same item) or from a different individual (i.e., a parent offering corrective feedback, the child disengaging from the task area, and the parent offering further corrective feedback). If a parent provided the same type of behavior for two items without a break in the middle (such as providing corrective feedback to item A and then again to item B with no other behavior occurring), then two separate instances of that behavior would be recorded (one for item A and one for item B).

This manuscript focuses on three categories of behaviors: *parent involvement*, *child disengagement*, and *technology issues*. Each of these behavioral categories are described as follows with examples of coded, category-specific behaviors.

*Parent involvement* included behavioral instances of when the parent helped or supported their child during the test administration. This included *corrective verbal assistance/support* beyond intended research member prompting that helped guide their child to the intended answer while answering a task item (e.g., “Remember what happened in the story? And then what happened after that?”); *general/vague verbal assistance/support* that does not hint at the intended answer (e.g., “Let’s tell them your answer so they can hear it”); *corrective feedback* to a child’s incorrect answer (i.e., “Are you sure that’s your answer? Is that the sound that the letter *a* makes?”); *general/vague verbal encouragement* to the child about their performance (e.g., “You’re doing great so far! Keep going!”); *physical assistance that impacts child performance* (e.g., parent physically writes a word for a child or helps the child spell a written word); *physical assistance to support task completion* (e.g., parent physically shows a piece of writing to the webcam because the child is unable to do so); and *verbal promise of a reinforcer* that might be delayed or immediate based on task engagement (e.g., “Finish this task and you’ll get to draw!” and “We can go to outside and play after you’re done with these activities”).

*Child disengagement* included behavioral instances when children lost focus during the test session or removed themselves from the testing environment. This included *attention disengagement from the immediate task* where the child appeared to visually disengage from the screen and the immediate task but physically remains in the task area; *physical disengagement from the task area* where the child physically left the immediate task area

**Table 2.** Frequencies, percentages (total and within categories), interrater reliability, and behavior representation across participants and assessment subtests.

Categories and coded behaviors	Frequency	% of total frequency	% of within-category frequency	Gwet's AC <sub>1</sub> [95% CI]	Behavior representation across participants		Behavior representation across assessment subtests	
					Descriptives (frequency) <sup>a</sup>	Breakdown by participant (% [frequency])	Descriptives (frequency) <sup>b</sup>	Breakdown by subtest (% [frequency])
<b>Parent involvement</b>	<b>456</b>	<b>68.37%</b>		<b>0.93</b> [0.91, 0.96]				
Corrective verbal assistance/support	104	15.59%	22.81%	0.96 [0.91, 1.00]	Min: 8.00 SD: 21.98 Med: 0.00 Mode: 0.00 Range: 0–80 Two parents accounted for 90% (13% [14] and 77% [80]). Remaining parents accounted for between 0% and 5% [0–5] each.	Min: 11.56 SD: 20.22 Med: 4.00 Mode: 4.00 Range: 0–64 Behaviors occurred during WJ IV VWS (62% [64]), WJ IV SWF (15% [16]), CELF-5 WS (7% [7]), EVT-3 (5% [5]), CELF-5 USP (4% [4]), WJ IV LWI (4% [4]), and WJ IV WA (4% [4]).		
General/vague verbal assistance/support	99	14.84%	21.71%	0.71 [0.44, 0.98]	Min: 7.62 SD: 10.67 Med: 2.00 Mode: 0.00 Range: 0–33 Four parents accounted for 87% (16% [16], 18% [18], 20% [20], and 33% [33]). Remaining parents accounted for between 0% and 6% [0–6] each.	Min: 11.00 SD: 8.76 Med: 8.00 Mode: N/A Range: 2–30 Behaviors occurred during EVT-3 (30% [30]), WJ IV WS (20% [20]), WJ IV LWI (12% [12]), NEPSY-II NM (9% [9]), WJ IV SWF (8% [8]), CELF-5 USP (7% [7]), WJ IV WA (6% [6]), CTOPP-2 RDN (5% [5]), and CELF-5 WS (2% [2]).		
Corrective feedback	77	11.54%	16.89%	0.87 [0.77, 0.97]	Min: 5.92 SD: 12.10 Med: 1.00 Mode: 0.00 Range: 0–41 Two parents accounted for 82% (29% [22] and 53% [41]). Remaining parents accounted for between 0% and 6% [0–5] each.	Min: 8.56 SD: 8.97 Med: 5.00 Mode: 5.00 Range: 1–30 Behaviors occurred during EVT-3 (39% [30]), WJ IV WS (19% [15]), CELF-5 WS (9% [7]), WJ IV LWI (9% [7]), WJ IV WA (6% [5]), WJ IV SWF (6% [5]), CTOPP-2 RDN (3% [2]), and NEPSY-II NM (1% [1]).		
General/vague verbal encouragement	114	17.09%	25.00%	0.98 [0.94, 1.00]	Min: 8.77 SD: 19.69 Med: 4.00 Mode: 0.00 Range: 0–73 Two parents accounted for 75% (11% [12] and 64% [73]). Remaining parents accounted for between 0% and 6% [0–7] each.	Min: 12.67 SD: 17.55 Med: 5.00 Mode: 3.00 Range: 0–55 Behaviors occurred during WJ IV VWS (48% [55]), WJ IV SWF (21% [24]), WJ IV LWI (12% [14]), EVT-3 (7% [8]), CTOPP-2 RDN (4% [5]), CELF-5 USP (3% [3]), CELF-5 WS (3% [3]), and WJ IV WA (2% [2]).		

(continued)

Table 2. Continued.

Categories and coded behaviors	Frequency	% of total frequency	% of within-category frequency	Gwet's AC <sub>1</sub> [95% CI]	Behavior representation across participants		Behavior representation across assessment subtests	
					Descriptives (frequency) <sup>a</sup>	Breakdown by participant (% [frequency])	Descriptives (frequency) <sup>b</sup>	Breakdown by subtest (% [frequency])
Verbal promise of a reinforcer	4	0.60%	0.88%	1.00 [1.00, 1.00]	Mn: 0.31 SD: 0.85 Med: 0.00 Mode: 0.00 Range: 0–3	Two parents accounted for 100% (25% [1] and 75% [3]).	Mn: 0.44 SD: 0.73 Med: 0.00 Mode: 0.00 Range: 0–2	Behaviors occurred during WJ IV SWF (50% [2]), NEPSY-II NM (25% [1]), and CELF-5 WS (25% [1]).
Physical assistance to support task completion	54	8.10%	11.84%	1.00 [1.00, 1.00]	Mn: 4.15 SD: 6.04 Med: 1.00 Mode: 0.00 Range: 0–17	Four parents accounted for 90% (13% [7], 22% [12], 24% [13], and 31% [17]). Remaining parents accounted for between 0% and 6% [0–3] each.	Mn: 6.00 SD: 12.17 Med: 0.00 Mode: 0.00 Range: 0–32	Behaviors occurred during WJ IV WS (59% [32]) and WJ IV SWF (41% [22]).
Physical assistance that impacts child performance	4	0.60%	0.88%	1.00 [1.00, 1.00]	Mn: 0.31 SD: 1.11 Med: 0.00 Mode: 0.00 Range: 0–4	One parent accounted for 100% [4].	Mn: 0.44 SD: 1.33 Med: 0.00 Mode: 0.00 Range: 0–4	Behaviors occurred during WJ IV SWF (100% [4]).
<b>Child disengagement</b>	<b>185</b>	<b>27.74%</b>		<b>0.86 [0.78, 0.95]</b>				
Attention disengagement from immediate task area	92	13.79%	49.73%	0.83 [0.69, 0.99]	Mn: 7.08 SD: 5.42 Med: 8.00 Mode: 2.00 Range: 0–19	Five children accounted for 67% (10% [9], 10% [9], 11% [10], 15% [14], and 21% [19]). Remaining children accounted for between 0% and 9% [0–8] each.	Mn: 10.22 SD: 7.24 Med: 8.00 Mode: 4.00 Range: 4–27	Behaviors occurred during WJ IV SWF (29% [27]), EVT-3 (16% [15]), WJ IV WS (13% [12]), WJ IV LWI (10% [9]), NEPSY-II NM (9% [8]), CELF-5 WS (8% [7]), CELF-5 USP (7% [6]), WJ IV WA (4% [4]), and CTOPP-2 RDN (4% [4]).
Physical disengagement from the task area	12	1.80%	6.49%	1.00 [1.00, 1.00]	Mn: 0.92 SD: 1.75 Med: 0.00 Mode: 0.00 Range: 0–6	Two children accounted for 75% (25% [3] and 50% [6]). Three children accounted for the remaining 25% (8% [1] each).	Mn: 1.33 SD: 1.32 Med: 1.00 Mode: 0.00 Range: 0–4	Behaviors occurred during NEPSY-II NM (33% [4]), EVT-3 (17% [2]), CELF-5 USP (17% [2]), WJ IV SWF (17% [2]), WJ IV WA (8% [1]), and WJ IV WS (8% [1]).
Expression of frustration or fatigue	68	10.19%	36.76%	0.88 [0.77, 0.98]	Mn: 5.23 SD: 6.46 Med: 2.00	Three children accounted for 67% (13% [9], 25% [17], and 29% [20]).	Mn: 7.56 SD: 5.00 Med: 7.00	Behaviors occurred during EVT-3 (22% [15]), WJ IV SWF (19% [13]), WJ IV WS

(continued)

**Table 2.** Continued.

Categories and coded behaviors	Frequency	% of total frequency	% of within-category frequency	Gwet's AC <sub>1</sub> [95% CI]	Behavior representation across participants		Behavior representation across assessment subtests	
					Descriptives (frequency) <sup>a</sup>	Breakdown by participant (% [frequency])	Descriptives (frequency) <sup>b</sup>	Breakdown by subtest (% [frequency])
Behavioral or attention challenge during task instructions	13	1.95%	7.03%	1.00 [1.00, 1.00]	Mode: 2.00 Range: 0–20  Mn: 1.00 SD: 1.53 Med: 0.00 Mode: 0.00 Range: 0–5	Remaining children accounted for between 0 and 9% [0–6] each.  Three children accounted for 76% (15% [2], 23% [3], and 38% [5]). Three children accounted for the remaining 24% (8% [1] each).	Mode: 2.00 Range: 2–15  Mn: 1.44 SD: 2.01 Med: 0.00 Mode: 0.00 Range: 0–5  Behaviors occurred during EVT-3 (38% [5]), NEPSY-II NM (31% [4]), CELF-5 USP (23% [3]), and CELF-5 VWS (8% [1]).	(19% [13]), WJ IV LWI (10% [7]), CELF-5 WS (10% [7]), NEPSY-II NM (9% [6]), WJ IV WA (4% [3]), CELF-5 USP (3% [2]), and CTOPP-2 RDN (3% [2]).
<b>Technology issues</b>	<b>26</b>	<b>3.90%</b>		<b>1.00</b> [1.00, 1.00]				
Assessor-initiated request to repeat answer	22	3.30%	84.62%	1.00 [1.00, 1.00]	Mn: 1.69 SD: 5.23 Med: 0.00 Mode: 0.00 Range: 0–19	One child accounted for 86% [19]. Two children accounted for the remaining 14% (5% [1] and 9% [2]).	Mn: 2.44 SD: 3.61 Med: 1.00 Mode: 0.00 Range: 0–11	Behaviors occurred during EVT-3 (50% [11]), CELF-5 WS (18% [4]), NEPSY-II NM (18% [4]), CTOPP-2 RDN (9% [2]), and CELF-5 USP (5% [1]).
Child-initiated request to repeat question	4	0.60%	15.38%	1.00 [1.00, 1.00]	Mn: 0.31 SD: 0.85 Med: 0.00 Mode: 0.00 Range: 0–3	Two children accounted for 100% (25% [1] and 75% [3]).	Mn: 0.44 SD: 1.01 Med: 0.00 Mode: 0.00 Range: 0–3	Behaviors occurred during EVT-3 (75% [3]) and CELF-5 USP (25% [1]).

Note. EVT-3 = Expressive Vocabulary Test, Third Edition; CELF-5 = Clinical Evaluation of Language Fundamentals, Fifth Edition (WS = word structure; USP = understanding spoken paragraphs); CTOPP-2 = Comprehensive Test of Phonological Processing, Second Edition (RDN = rapid digit naming); NEPSY-II = A Developmental Neuropsychological Assessment, Second Edition (NM = narrative memory); WJ IV = Woodcock-Johnson IV (SWF = sentence writing fluency, WS = writing samples); CI = confidence interval (95%); GWET's AC<sub>1</sub> = Gwet agreement coefficient for first-order chance correction; Mn = mean; SD = standard deviation; Med = median. Due to rounding, percentages may not equal 100%.  
<sup>a</sup>Refers to the descriptive statistics for the frequency of the specified behavior across all participants (regardless of assessment subtest).  
<sup>b</sup>Refers to the descriptive statistics for the frequency of the specified behavior across all assessment subtests (regardless of participant).

(i.e., moving from the area captured by their Zoom video either to another side of the room or to a different room entirely); *expression of frustration or fatigue* with the task, including verbal and physical indications of anger, boredom, or lethargy; and *behavioral or attention challenge during task instructions* where the child appeared unable to hear task instructions or particular items due to behavior or attention dysregulation.

*Technology issues* included two instances where a technology-related problem occurred with engaging in the task. This included *assessor-initiated request to repeat answer* where the research team member noticed a technology-related error (audio, microphone, Internet lag, etc.) and repeated either task instructions or a stimulus item and *child-initiated request to repeat question* where the child requested the research team member to repeat the task instructions or a stimulus item due to a technology-related error.

### Assessment procedures

Each participant completed two assessment sessions: Session A (oral and reading) and Session B (writing). Each session lasted about an hour and a half (however, the discussed writing assessments took approximately 30 min during Session B). Session B took place one to two weeks after Session A. Four children required additional time for Session A, thus an additional session was scheduled. The first and second authors administered Session A assessments, and the third author administered Session B assessments. Authors include two postdoctoral researchers and a junior faculty member, all with extensive experience administering study assessments to school-age children with ASD in face-to-face settings. Parents chose the most appropriate location in their homes for the sessions (e.g., kitchen table, office space). All assessments were video recorded via Zoom so that the assessment could later be double scored by another research team member. Parents were permitted to be present during the testing process but were instructed not to provide answers or assist children in answering questions. During assessment administration, eight (61.54%) of the children had a parent or caregiver present for all or most of the two sessions. The same children had a parent present during both Session A and Session B. Individual assessments were discontinued or rescheduled if a child exhibited fatigue, three times across Session A, or due to non-response from parents, one instance across Session B, and reward systems were incorporated to minimize participant discomfort (for three children).

### Standardized measures

All participants were assessed with standardized oral language, reading, and writing assessments using an adapted

tele-assessment approach via Zoom. All assessments are clinical tools which have been standardized for face-to-face administration. Figure 1 highlights overall and participant-level performance across included assessment subtests. Oral language assessments included Word Structure and Understanding Spoken Paragraphs from the Clinical Evaluation of Language Fundamentals, Fifth Edition (CELF-5; Wiig et al., 2013); the Expressive Vocabulary Test, Third Edition (EVT-3; Williams, 2018); and the Narrative Memory subtest from the NEPSY-II (Korkman et al., 2007). Reading assessments included Rapid Digit Naming from the Comprehensive Test of Phonological Processing, Second Edition (CTOPP-2; Wagner et al., 2013) and Letter-Word Identification and Word Attack from the Woodcock-Johnson Tests of Achievement, Fourth Edition (WJ IV; Schrank et al., 2014). Writing assessments included Sentence Writing Fluency and Writing Samples (Form A) from the WJ IV. See Supplemental Figure 1 for full assessment information, including subtest descriptions, psychometric evidence, adaptations made, and estimated administration times.

Subtests from the CELF-5 and EVT-3 were administered via Q-Global, an online administration platform created by Pearson Education, Inc. for electronic assessment delivery. The remainder of the assessments were adapted for online administration using emerging best practices and practical considerations to mimic in-person administration guidelines using Microsoft PowerPoint ( $n = 5$ ) or were not adapted as they required no visual stimuli ( $n = 2$ ). The available validity evidence to date for other neuropsychological assessments does not suggest that validity would be threatened by changing mode of administration, but this evidence does not cover all types of assessments nor all clinical conditions (Brearly et al., 2017; Sutherland et al., 2017).

### Oral language assessments

To administer the structural language assessments from the EVT-3 and CELF-5 (Word Structure), an electronic stimulus book was accessed by the assessor via Q-Global and shared on a screen for the child during the assessment sessions. The assessors scored child responses using a paper copy of the score form during the session. For higher-order language assessments including NEPSY-II Narrative Memory and CELF-5 Understanding Spoken Paragraphs, no adaptations were made, as these assessments are administered orally with no visuals.

### Reading assessments

Test items from the Letter-Word Identification and Word Attack subtests from the WJ IV and the Rapid Digit Naming subtest from the CTOPP-2 were administered via Microsoft PowerPoint. Test items and accompanying



images were transferred to PowerPoint slides, which were then shared via Zoom during the testing sessions. For items on the WJ IV that required a pointing response, the assessor allowed the children to assume cursor control, granting permission for the participant to use the computer mouse. The assessors created a “practice item” where the children could rehearse taking control of the cursor and gesturing to their answer; this practice item was incorporated into the slides. For children who were unable to take control of the mouse, a form of the PowerPoint slides was given where all possible responses on a test item were surrounded by a colored border. For example, if the response options included letters *P*, *F*, *G*, and *J*, each letter would have a different colored border surrounding it. Before answering the test item, children were asked to name every color shown on that slide. All children were able to name these colors, and this response option was used by four students. Then, the assessor asked the child to identify which color was around the correct response. Children responded verbally to all other WJ IV items that did not require pointing and all items from the CTOPP-2 subtest. The assessor marked child responses on paper forms.

### Writing assessments

Both WJ IV writing subtests were transferred to Microsoft PowerPoint slides and designed to visually emulate the student response booklet. Participating children were asked to either write on blank paper with a pen or pencil or on a whiteboard with a marker and to physically hold their responses up to their webcam when done for the assessor to see their work. For sentence writing fluency, each practice item was shown on individual Microsoft PowerPoint presentation slides, and animations of black rectangles were animated sequentially to highlight where the child should focus their attention during the instructions (i.e., the three words, the provided stimulus picture, and the blank space for writing). Test items did not use the guiding box animations and were presented one by one in full (i.e., the three words, the provided stimulus picture, and the blank space for writing). The timer was paused in-between test items for the duration that it took the child to visually show their response to the webcam. Once the next item was shown to the participant, the timer resumed. For writing samples, as no practice items are provided, testing began with the first item and proceeded sequentially until the child hit their ceiling. Items were transferred into Microsoft PowerPoint from the stimulus book and were shown to children one at a time.

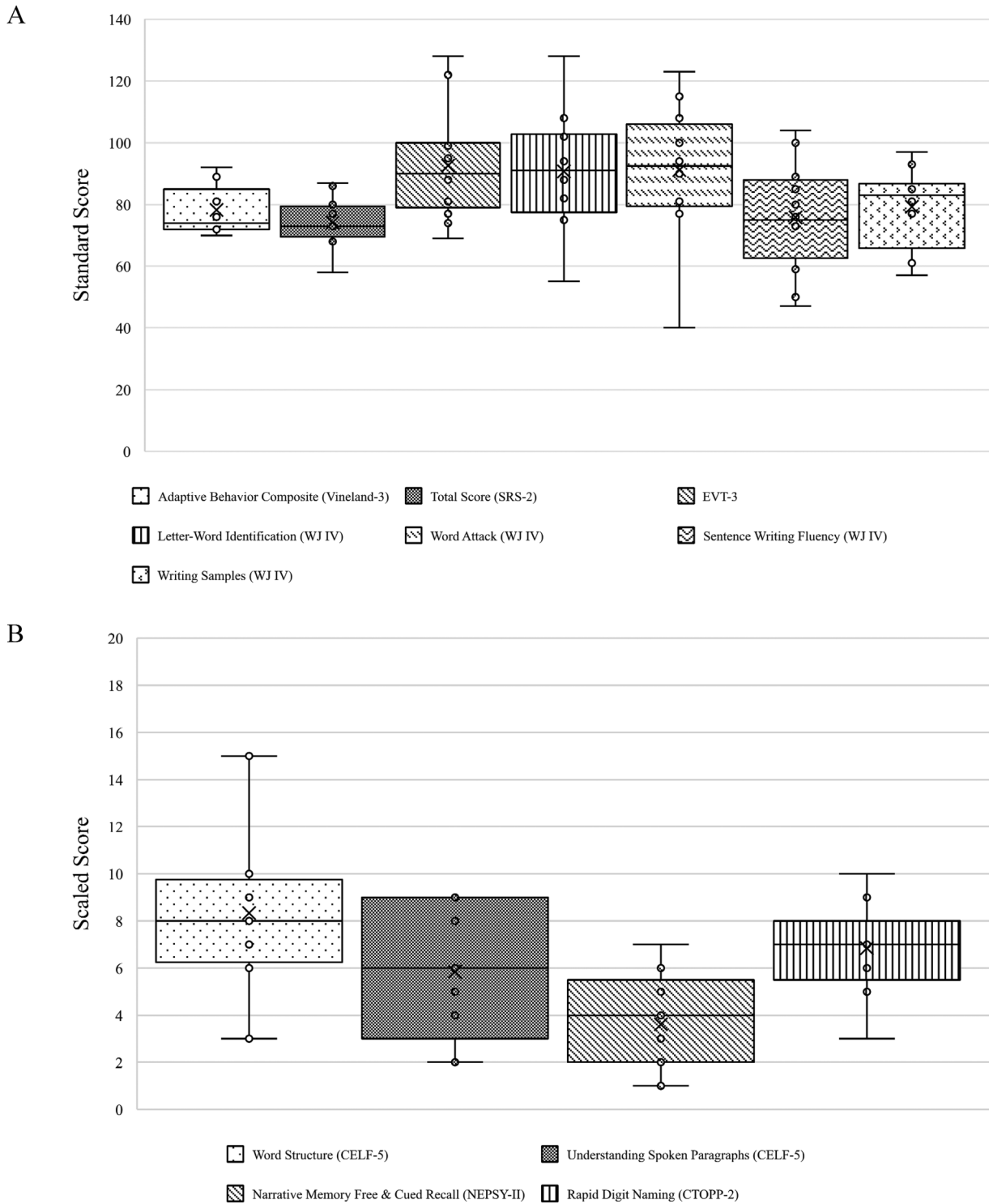
## Results

Table 2 summarizes the behavior frequencies and percentages (overall and within category) and behavior

representation across participants and assessment subtests. Parent involvement behaviors made up approximately two-thirds of the coded behaviors (68.37%, frequency [F] = 456). The most frequent within-category behaviors included general/vague verbal encouragement (25.00%, F = 114), corrective verbal assistance/support (22.81%, F = 104), general/vague verbal assistance/support (21.71%, F = 99), corrective feedback (16.89%, F = 77), and physical assistance to support task completion (11.84%, F = 54). There were few instances of physical assistance that impacts child performance (0.88%, F = 4) and verbal promise of a reinforcer (0.88%, F = 4). Across participant sessions, parent involvement appeared to split into two smaller subsets of parents: (1) between one and four parents per behavior who accounted for most (75%) of the occurrences, and (2) remaining parents who accounted for few if any of the occurrences (<10%). Parent involvement behaviors generally occurred across most assessment subtests but at variable rates, with each behavior showing elevated frequencies on a small number of subtests (though with the specific assessment subtests varying by behavior; see Table 2).

*Child disengagement* behaviors made up approximately one-fourth of the coded behaviors (27.74%, F = 185). The most frequent within-category behaviors were *attention disengagement from the immediate task* (49.73%, F = 92) and *expression of frustration or fatigue* (36.76%, F = 68). In contrast, there were fewer instances of *physical disengagement from the task area* (6.49%, F = 12) and *behavioral or attention challenge during task instructions* (7.03%, F = 13). Across sessions, child disengagement behaviors appeared to occur more frequently across many different participants (as indicated by relatively lower frequencies but greater coverage across children). Some assessment subtests were associated with consistently higher rates of disengagement behaviors relative to other subtests, particularly the EVT-3 (16%–38%) and the NEPSY-II Narrative Memory subtest (9%–33%).

*Technology issues* behaviors occurred overall infrequently compared to the other categories (3.90%, F = 26). Within-category behaviors skewed towards *assessor-initiated request to repeat answer* in comparison to *child-initiated request to repeat question* (84.62%, F = 22 and 15.38%, F = 4, respectively). Across participants, technology issues for the assessor occurred with a small number of participants, with one participant showing elevated problems compared to the other two (86% compared to 5% and 9%). Similarly, child-oriented technology issues were isolated to two children, with more relative issues for one child (75%). Across subtests, assessor-oriented technology problems seemed to primarily occur during the EVT-3 (50%). A similar trend was observed for the child-oriented technology problems with the EVT-3 showing the highest number of instances (75%).



**Figure 1.** Boxplots of (A) participant standard and *T* scores and (B) scaled scores for included study assessments. *Note.* SRS-2 = Social Responsiveness Scale, Second Edition; EVT-3 = Expressive Vocabulary Test, Third Edition; WJ IV = Woodcock-Johnson Tests of Achievement, Fourth Edition; CELF-5 = Clinical Evaluation of Language Fundamentals, Fifth Edition; NEPSY-II = A Developmental NEuroPSYchological Assessment, Second Edition; CTOPP-2 = Comprehensive Test of Phonological Processing, Second Edition. Vineland-3 scores are based off the Parent/Caregiver Domain-Level Form. SRS-2 scores are based off the Total Score from the Parent School-Age Form. A horizontal line represents the median score. A cross represents the mean. In A, all scores are standard scores except for SRS-2 (*T* score).

## Relationships between behaviors, autism symptomatology, and adaptive behavior skills

We examined relationships between behavior frequencies (across all assessments) and standardized measures of autism symptom severity and adaptive behavior skills (Table 3). Given limited sample sizes, all correlation coefficients were calculated using Kendall's Tau-b. We first examined relationships for the reading and writing assessment sessions separately before also examining relationships for both sessions combined. For parent involvement, the SRS-2 showed a positive, strong relationship with corrective verbal assistance/support during Session A and corrective feedback during Session B; neither were statistically significant after combining the sessions. For child disengagement, the SRS-2 showed a negative, moderate relationship with behavioral or attention challenge during task instruction during Session A; this remained significant for the combined session given that behaviors were observed during Session B. The Vineland-3 ABC showed a positive, strong relationship with behavioral or attention challenge during task instruction and a positive, moderate relationship with physical disengagement from the task area during Session A. Combining the sessions provided no additional information given that none of the aforementioned behaviors occurred during Session B. For technology issues as well as for the remaining correlation coefficients for parent involvement and child disengagement, coefficients varied from weak to moderate but were not statistically significant.

## Discussion

This study sought to examine the feasibility of conducting tele-assessment oral language, reading, and writing assessments with a small group of children diagnosed with ASD. A primary goal of the current study was to examine to what extent the presence of a parent during such assessments may have impacted data collection in the tele-assessment format. We first present information related *parent involvement*, then *child disengagement*, followed by *technology issues*. We also include limitations and implications for future oral language, reading, and writing assessments delivered over Zoom (or other platforms) to children with ASD.

### Parent involvement

Parent involvement can be categorized into assistance and encouragement that (1) does not lead the child to the correct answer and (2) assistance that leads the child to the correct response. Non-leading assistance included encouraging comments, repetition of task instructions, or physical assistance in holding up a whiteboard or paper to

the webcam. Overall, these types of non-leading behaviors accounted for 59.07% of all parent involvement codes and appeared more evenly were spread out across assessments and parents (Table 2). This was not the case for leading assistance behaviors, which were concentrated among two individual caregivers. In many instances, the child would first provide an incorrect response before the parent interjected or began to provide physical assistance in the task (when writing). One caregiver (an applied behavior analyst therapist who worked with the child at home several days a week) accounted for 14.29% ( $n=2$ ) of these leading assistance codes, and another caregiver accounted for 85.71% ( $n=12$ ). This second caregiver was the mother of a 7-year-old girl who had low-average adaptive behavior skills and severe difficulty with social interaction (as measured by the Vineland-3 and SRS-2, respectively). It is possible that these caregivers were more involved with leading behaviors due to these children's struggles with social interactions and lower adaptive behavior skills, as compared to the other children in the sample. However, other participants who did score similarly on these parent reports did not show high levels of leading parent assistance, suggesting that these behaviors were specific to these caregiver-child dyads. This suggests that parental assistance behaviors do not depend solely on the abilities of the child and may depend on the attitude of the parent/caregiver and their relationship with their child.

An interesting difference emerged when looking at corrective assistance compared to corrective feedback in relation to the level of autism symptomatology during the two assessment sessions. Findings highlight strong relationships between autism symptomatology and corrective verbal assistance (only during the reading and oral language assessment sessions) as well as corrective feedback (only during the writing assessment session), both of which only held for specific assessment sessions (Table 3). These noted differences may be due to differences in the assessment demands as well as general behavior frequency, given that two parents accounted for most of the behaviors for both types of corrective behaviors. However, these initial relationships raise important considerations that the presence of parent behaviors may not only be related to the level of autism symptomatology of the child being assessed but also to the type of assessment being administered (and when a parent might provide assistance or feedback to a child's response).

Non-leading assistance is unlikely to impact the validity of assessment scores (and was largely unrelated to child's level of autism symptomatology or adaptive behavior skills; see Table 3), but leading assistance may be impactful to the validity of the child's assessment results. To protect against this threat to validity, assessors marked items as "incorrect" when the child first provided an incorrect response but later changed their response due to assistance

from a parent. However, caregiver-provided assistance or feedback could potentially impact performance on the current item, on later items of that assessment, and on other administered assessments. In the current study, we observed that two caregivers accounted for most of the corrective assistance and feedback provided, suggesting this to be more problematic for individual cases rather than entire samples of children. Therefore, administrators of adapted tele-assessments should be trained to recognize leading behaviors exhibited by caregivers who are present during the assessment and address them accordingly to avoid potentially diminishing the validity of assessment results (particularly for individual cases). We were not able to measure the threat of early assistance and its potential impact on later test items given: the use of standardized assessments with different numbers of items administered due to different basals and ceilings, the lack of guidelines for altering scoring procedures for such guidance, and the current study design focused more on understanding the tele-assessment context as a whole. We recognize that future studies incorporating tele-assessment practices should focus on leading assistance behaviors and if this type of support directly impacts performance on assessments.

A final important takeaway is that not all assessment sessions had a parent present. Out of the 13 included participants, only 8 children (61.54%) had a parent or other caregiver remain with them for most of the assessment session durations. For children without a parent or other caregiver present, they needed to navigate the tele-assessment context themselves. This includes the need to self-regulate their engagement during the session. We were fortunate that no child refused to join our assessment sessions or close Zoom unexpectedly. As seen in Table 2, many parents joined solely to serve a support role to help their child engage with the assessor without influencing the task outcomes. Therefore, the presence of a parent or other caregiver does not immediately disqualify the findings from tele-assessment. However, as noted above, a small number of parents provided consistent indirect support to their child in the form of leading behaviors that might have influenced child performance. Taken altogether, these observations highlight the importance of not only noting if a parent or other caregiver was present during a tele-assessment session, but to what extent the parent interacted or engaged with their child during the assessments. The presence of a parent providing only general support and encouragement may be interpreted as having a similar effect on child performance as to having no parent involved whatsoever (as the parent fulfilled a role similar to that of an assessor during in-person assessment administration), but both of those contexts are quite different from a parent or caregiver who is involved and engaging in ways that might influence their child's performance.

### *Child disengagement*

Most of the identified disengagement behaviors (49.73%, or 92 instances overall) included instances where the child was distracted or not paying attention to the task instructions but remained in the task area. When children left the task area, which did not happen often (only 6.49% of coded disengagement behaviors, or 12 instances overall), the assessor could call out to the child and quickly redirect them. One child showed extreme resistance (50% of coded behaviors, or six instances overall) and was difficult to persuade to return to the computer area and complete the assessment battery. To combat this, the assessor provided breaks to minimize testing fatigue. While some children expressed outright frustration with the length or difficulty of the tasks (36.76% of coded disengagement behaviors, or 68 instances overall), assessors were usually able to redirect these children with encouraging comments or assurances that the task was almost complete.

Relationships between disengagement behavior frequencies and levels of autism symptomatology and adaptive behavior skills varied. Both disengaging visually from the immediate task area as well as expressing frustration or fatigue appeared overall unrelated to either autism symptomatology or adaptive behavior skills. In contrast, physically disengaging appeared to be positively related to adaptive behavior skills while demonstrating regulation challenges appeared to be negatively related to autism symptomatology and positively related to adaptive behavior skills. Given the overall low frequency of these behaviors (12 and 13 instances, respectively) across a small number of children (two and three children accounted for most of the behaviors, respectively) and the restricted ranges for both measures of autism symptomatology and adaptive behavior skills (Table 2), we interpret these findings with caution. In general, given the low number of instances for these behaviors, overall takeaways appear similar to those offered for parent involvement behaviors: researchers and educators should be mindful of individual differences in children with autism, which might influence the ability to make valid interpretations about assessment performance.

### *Technology issues*

We included two codes related to technology issues: (1) instances of the assessor asking the child to repeat an answer due to a technology issue (i.e., lagging internet, muting the computer volume, background noise, etc.) and (2) instances of the child requesting directional clarification (due to technology issues). The assessors rarely asked the children to repeat their answer due to a technology issue (only 23 instances across all codes, or 3.30%), and the majority of the requests (86%) were within the assessment sessions for one child with a weak internet connection (relative to other participants). Children rarely requested

**Table 3.** Kendall's Tau-b non-parametric correlation coefficients between parent involvement, child disengagement, and technology issues and standardized measures of autism symptom severity and adaptive behavior skills across assessment sessions.

	Assessment session A—Reading and oral language		Assessment session B— Writing		Assessment sessions combined	
	SRS-2	ABC	SRS-2	ABC	SRS-2	ABC
<b>Parent involvement<sup>a</sup></b>						
Corrective verbal assistance/support	0.64*	−0.36	0.33	0.12	0.55	−0.12
General/vague verbal assistance/support	0.39	0.08	0.11	0.42	0.40	0.11
Corrective feedback	0.35	−0.04	0.71*	−0.20	0.38	−0.04
General/vague verbal encouragement	0.29	0.21	−0.04	0.57	0.12	0.43
Verbal promise of a reinforcer	0.22	−0.36	0.22	−0.36	0.22	−0.36
Physical assistance to support task completion	− <sup>c</sup>	− <sup>c</sup>	0.44	0.00	0.44	0.00
Physical assistance that impacts child performance	− <sup>c</sup>	− <sup>c</sup>	0.51	0.22	0.51	0.22
<b>Child disengagement<sup>b</sup></b>						
Attention disengagement from immediate task area	0.03	0.14	0.10	0.23	0.01	0.16
Physical disengagement from the task area	−0.37	0.58*	0.17	0.00	−0.27	0.50*
Expression of frustration or fatigue	0.35	−0.04	0.23	0.18	0.35	0.06
Behavioral or attention challenge during task instructions	−0.50*	0.70**	− <sup>c</sup>	− <sup>c</sup>	−0.50*	0.70**
<b>Technology issues<sup>b</sup></b>						
Assessor-initiated request to repeat answer	−0.22	−0.08	− <sup>c</sup>	− <sup>c</sup>	−0.22	−0.08
Child-initiated request to repeat question	−0.41	0.39	− <sup>c</sup>	− <sup>c</sup>	−0.41	0.39

Note. \* $p < .05$ , \*\* $p < .01$ . SRS-2 = Total Score, Social Responsiveness Scale, 2nd Edition; ABC = Adaptive Behavior Composite, Vineland Adaptive Behavior Scales, 3<sup>rd</sup> Edition.

$n = 8$ . The same children had a parent or other adult present for most of the time during the reading and writing assessment sessions.

$n = 13$ . All participants were included in these analyses regardless of if a parent were present or not.

<sup>c</sup>No correlation coefficient could be calculated because the behavior frequencies were zero across participants.

clarification due to technology issues (four instances across all codes, or 0.60%), as a majority of the children appeared to understand the assessor and did not ask for repeated questions or directions. There also did not appear to be any significant group-level relationships between the number of technology issues and the level of autism symptomatology or the adaptive behavior skills of the child. This was expected given that the noted technology problems might be more related to issues with the technology setup rather than to a child's behavior. Altogether, these results are promising that technology problems specific to responding to assessment items seem to be relatively uncommon.

### Limitations

Our study was not exhaustive to all the potential technology challenges faced during the pandemic; as such, while this small-scale effort proved feasible, studies involving larger numbers of families may introduce new problems not observed here. Additionally, the order of assessments was not randomized, potentially impacting child performance. Other limitations include disruptions (e.g., background noise, lagging internet) that did not frequently occur during our data collection but have the potential to make it difficult for children to complete tasks or follow

directions. It is possible that assessment adaptations may impact the validity of these results. However, for the subtests in which Microsoft PowerPoint was used, the research team designed the tele-assessment layout to mimic the testing layout provided during face-to-face assessment. Improvements that could improve the validity of online administered assessments could include electronic testing formats that can be accessed online and presented to the test taker via a shared screen.

Additionally, the data regarding parental involvement is concentrated around a subset of parents who provided the leading and assistance type of behaviors. Of the eight parents and caregivers who were present during the assessment sessions, a majority of the codable behaviors that have the potential to impact assessment scores (i.e., verbal help to the child that hints at the answer) came from a small number of caregivers. The two children involved in these interactions had adaptive behavior scores on the lower end of the range of this sample; however, several other children in the sample had scores within a few points of these two children, so it is likely that these behaviors are attributable to the parent rather than the child. Future research could attempt to establish information on parenting style, which could possibly relate to involvement. Finally, this sample was limited to children aged 6 to 9, and those

who scored above 70 on measures of verbal IQ and adaptive behavior, which likely impacts the generalizability of the findings for children with ASD who have intellectual disability (ID). It may be that children with ASD and ID are more likely to disengage from testing sessions, and further research should examine the reliability of reading and writing tele-assessments for this population, as well as children who are in the preschool or kindergarten age range.

### Implications

This research supports the feasibility of adapted oral language, reading, and writing tele-assessments for children with ASD. Overall, the parents of participants in this study did not provide many instances of leading assistance that could potentially impact their child's assessment scores. Additionally, instances of children becoming disengaged were not frequent, and when children did become disengaged, the assessor was often able to redirect the child, at times with a simple verbal reminder and at other times with a timer for a planned break. Combined with the low incidence of technology-related issues, these data suggest that conducting online language, reading, and writing assessments is practical for early elementary-aged children with ASD who do not have ID. Further implementation may uncover additional issues and concerns not captured in the current study, but initial findings suggest that potential areas of concern (e.g., parental interference during testing) are infrequent and can be addressed easily by a trained research team.

While the data presented in this manuscript seem to support the use of tele-assessments for children with ASD, more needs to be done to ensure the validity and reliability of these assessments (Pearson Education, 2020; Riverside Insights, 2020; Wright, 2020). While some research supports levels of agreement between tele-assessment and in-person assessment scores (e.g., Sutherland et al., 2019), the requirements for psychometric equivalency set forth by the formal recommendations for educational and psychological testing have not yet been met (Krach et al., 2020). Different types of assessments also have different demands for the child depending on the test items; in turn, certain child characteristics and the extent of parent involvement may differ solely based on such demands. More research is needed that examines the psychometrics of adapted tele-assessments for children with ASD in all age ranges. This research should consider testing environments that are managed by the student or caregiver, not the facilitator, and supports the validity of results from adapted tele-assessment for this population (Pearson Education, 2020; Riverside Insights, 2020; Wright, 2020).

Moving forward, clear guidelines must be created for the adapting administration of oral language, reading, and writing assessments for tele-assessment delivery for children with developmental disabilities such as ASD.

Guidelines should consider the variance of ASD symptomology across test takers. Guidelines on adapted tele-assessments should also include the tailoring of stimulus images for electronic presentation, modification of directions that cannot be repeated or could be easily misinterpreted, and alteration of questions that require pointing or other physical manipulations. Guidelines for caregivers who remain in the child's vicinity during the assessment should be created that include acceptable prompting or redirection and reiterate the importance of providing only behavior support. Furthermore, studies should describe the potential impacts that parents and caregivers may have on administered assessments and child outcomes. Additional guidelines on responding to technology issues during assessment sessions should also be explored.

### Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

### Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was supported by the University of Virginia Supporting Transformative Autism Research Initiative and the Postdoctoral research Fellow Training Program Grant from the National Center for Special Education research at the Institute of Education Sciences (R324B180034).

### ORCID iD

Carlin Conner  <https://orcid.org/0000-0002-9790-9047>

### Supplemental material

Supplemental material for this article is available online.

### References

- Baharav, E., & Reiser, C. (2010). Using telepractice in parent training in early autism. *Telemedicine and e-Health, 16*(6), 727–731. <https://doi.org/10.1089/tmj.2010.0029>
- Brearly, T. W., Shura, R. D., Martindale, S. L., Lazowski, R. A., Luxton, D. D., Shenal, B. V., & Rowland, J. A. (2017). Neuropsychological test administration by videoconference: A systematic review and meta-analysis. *Neuropsychology Review, 27*(2), 174–186. <https://doi.org/10.1007/s11065-017-9349-1>
- Constantino, J. N., & Gruber, C. P. (2012). *Social responsiveness scale: SRS-2* (p. 106). Western Psychological Services.
- Friard, O., Gamba, M., & Fitzjohn, R. (2016). BORIS: A free, versatile open-source event-logging software for video/audio coding and live observations. *Methods in Ecology and Evolution, 7*(11), 1325–1330. <https://doi.org/10.1111/2041-210X.12584>

- Gwet, K. L. (2014). *Handbook of interrater reliability: The definitive guide to measuring the extent of agreement among raters*. Advanced Analytics, LLC.
- Henry, A. R., & Solari, E. J. (2020). Targeting oral language and listening comprehension development for students with autism spectrum disorder: A school-based pilot study. *Journal of Autism and Developmental Disorders*, *50*(10), 3763–3776. <https://doi.org/10.1007/s10803-020-04434-2>
- Ketterlin-Geller, L. R. (2008). Testing students with special needs: A model for understanding the interaction between assessment and student characteristics in a universally designed environment. *Educational Measurement: Issues and Practice*, *27*(3), 3–16. <https://doi.org/10.1111/j.1745-3992.2008.00124.x>
- Klein, D. (2018). Implementing a general framework for assessing interrater agreement in Stata. *The Stata Journal: Promoting Communications on Statistics and Stata*, *18*(4), 871–901. <https://doi.org/10.1177/1536867X1801800408>
- Korkman, M., Kirk, U., & Kemp, S. (2007). *NEPSY-II: Clinical and interpretive manual*. The Psychological Corporation.
- Krach, S. K., Paskiewicz, T. L., & Monk, M. M. (2020). Testing our children when the world shuts down: Analyzing recommendations for adapted tele-assessment during COVID-19. *Journal of Psychoeducational Assessment*, *38*(8), 923–941. <https://doi.org/10.1177/0734282920962839>
- McIntyre, N. S., Solari, E. J., Grimm, R. P., Lerro, L. E., Gonzales, J. E., & Mundy, P. C. (2017). A comprehensive examination of reading heterogeneity in students with high functioning autism: Distinct reading profiles and their relation to autism symptom severity. *Journal of Autism and Developmental Disorders*, *47*(4), 1086–1101. <https://doi.org/10.1007/s10803-017-3029-0>
- Nation, K., Clarke, P., Wright, B., & Williams, C. (2006). Patterns of reading ability in children with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *36*(7), 911–919. <https://doi.org/10.1007/s10803-006-0130-1>
- Pearson Education. (2020). Telepractice norms and validity of performance-based tests [Online Resource Document]. <https://www.pearsonassessments.com/content/dam/school/global/clinical/us/assets/telepractice/telepractice-norms-and-validity-of-performance-based-tests.pdf>
- Pennington, B. F., McGrath, L. M., & Peterson, R. L. (2020). Autism spectrum disorder. In Pennington, B. F., McGrath, L. M., & Peterson, R. L. (Eds.), *Diagnosing learning disorders: From science to practice* (3rd ed., pp. 258–311). Guilford Press.
- Regina Molini-Avejonas, R. D., Rondon-Melo, S., De La Higuera Amato, C. A., & Samelli, A. G. (2015). A systematic review of the use of telehealth in speech, language and hearing sciences. *Journal of Telemedicine and Telecare*, *21*(7), 367–376. <https://doi.org/10.1177/1357633X15583215>
- Riverside Insights. (2020). Remote assistance guidelines [Customer Letter]. <https://cms.riversideinsights.com/uploads/735e4b05cf5142c695b3b4b5d3fb42a1.pdf>
- Schrank, F. A., Mather, N., & McGrew, K. S. (2014). *Woodcock-Johnson IV Tests of Achievement*. Riverside.
- Solari, E. J., Henry, A. R., McIntyre, N. S., Grimm, R. P., & Zajic, M. (2020). Testing the effects of a pilot listening comprehension and vocabulary intervention for individuals with autism. *Research in Autism Spectrum Disorders*, *71*, 101501. <https://doi.org/10.1016/j.rasd.2019.101501>
- Sparrow, S. S., Cicchetti, D. V., & Saulnier, C. A. (2016). *Vineland Adaptive Behavior Scales* (3rd ed.). Pearson.
- Sutherland, R., Trembath, D., Hodge, A., Drevensek, S., Lee, S., Silove, N., & Roberts, J. (2017). Telehealth language assessments using consumer grade equipment in rural and urban settings: Feasible, reliable and well tolerated. *Journal of Telemedicine and Telecare*, *23*(1), 106–115. <https://doi.org/10.1177/1357633X15623921>
- Sutherland, R., Trembath, D., Hodge, M. A., Rose, V., & Roberts, J. (2019). Telehealth and autism: Are telehealth language assessments reliable and feasible for children with autism? *International Journal of Language & Communication Disorders*, *54*(2), 281–291. <https://doi.org/10.1111/1460-6984.12440>
- Vismara, L. A., McCormick, C., Young, G. S., Nadhan, A., & Monlux, K. (2013). Preliminary findings of a telehealth approach to parent training in autism. *Journal of Autism and Developmental Disorders*, *43*(12), 2953–2969. <https://doi.org/10.1007/s10803-013-1841-8>
- Vismara, L. A., Young, G. S., & Rogers, S. J. (2012). Telehealth for expanding the reach of early autism training to parents. *Autism Research and Treatment*, *2012*. <https://doi.org/10.1155/2012/121878>
- Wagner, R. K., Torgesen, J. K., Rashotte, C. A., & Pearson, N. A. (2013). CTOPP-2: Comprehensive test of phonological processing: Pro-ed.
- Waite, M. C., Theodoros, D. G., Russell, T. G., & Cahill, L. M. (2010). Internet-based telehealth assessment of language using the CELF-4. *Language, Speech, and Hearing Services in Schools*, *41*(4), 445–458. [https://doi.org/10.1044/0161-1461\(2009/08-0131\)](https://doi.org/10.1044/0161-1461(2009/08-0131))
- White, L. C., Law, J. K., Daniels, A. M., Toroney, J., Vernoia, B., Xiao, S., Feliciano, P., & Chung, W. K. (2021). Brief report: Impact of COVID-19 on individuals with ASD and their caregivers: A perspective from the SPARK cohort. *Journal of Autism and Developmental Disorders*, *51*(10), 3766–3773. <https://doi.org/10.1007/s10803-020-04816-6>
- White, S. W., Stoppelbein, L., Scott, H., & Spain, D. (2021). It took a pandemic: Perspectives on impact, stress, and telehealth from caregivers of people with autism. *Research in Developmental Disabilities*, *113*, 103938. <https://doi.org/10.1016/j.ridd.2021.103938>
- Wiig, E. H., Semel, E., & Secord, W. A. (2013). *Clinical evaluation of language fundamentals—fifth edition (CELF-5)*. Pearson.
- Williams, K. T. (2018). *Expressive vocabulary test* (3rd ed.). NCS Pearson.
- Wright, A. J. (2020). Equivalence of remote, digital administration and traditional, in-person administration of the Wechsler Intelligence Scale for Children (WISC-V). *Psychological Assessment*, *32*(9), 809–817. <https://doi.org/10.1037/pas0000939>
- Zajic, M. C., & Wilson, S. E. (2020). Writing research involving children with autism spectrum disorder without a co-occurring intellectual disability: A systematic review using a language domains and mediational systems framework. *Research in Autism Spectrum Disorders*, *70*, 101471. <https://doi.org/10.1016/j.rasd.2019.101471>