UNIVERSITY OF CALIFORNIA
RIVERSIDE

Telehealth Autism Spectrum Disorder Evaluations: Efficacy and Caregiver Perceptions

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Education

by

Michelle Heyman

September 2023

Dissertation Committee:
Dr. Jan Blacher, Co-Chairperson
Dr. Katherine Meltzoff, Co-Chairperson
Dr. Stephanie Moore

The Dissertation of Michelle Heyman is approved:

_____

_____

Committee Co-Chairperson

_____

Committee Co-Chairperson

University of California, Riverside

ABSTRACT OF THE DISSERTATION

Telehealth Autism Spectrum Disorder Evaluations: Efficacy and Caregiver Perception

by

Michelle Heyman

Doctor of Philosophy, Graduate Program in Education
University of California, Riverside, September 2023
Dr. Jan Blacher, Co-Chairperson, & Dr. Katherine Meltzoff, Co-Chairperson

Autism spectrum disorder (ASD) is a heterogeneous disorder characterized by social communication deficits and the presence of repetitive and restrictive behaviors. Early ASD diagnosis results in earlier interventions and is linked to a better long-term prognosis. The gold standard assessment tool for ASD evaluations is the ADOS-2, an in-person evaluation that elicits specific behavioral responses to predefined activities. However, the COVID-19 pandemic led to a need for telehealth-based ASD assessments. Several ASD assessment tools were developed; however, many were designed to assess young children. This preliminary study aimed to evaluate the accuracy and effectiveness of two telehealth assessment tools, TELE-ASD-PEDS and TELE-ASD-KIDS, compared to the ADOS-2. Clinician diagnostic impressions were further analyzed to determine if child and family variables influenced diagnostic impressions and clinician diagnostic confidence levels at the telehealth appointment. Additionally, family experiences were collected after each appointment to understand the assessment experience. The participants were assessed twice, once through telehealth and once in-person (N = 27),

and were randomly assigned to one of the assessment styles first. The results showed promising outcomes for telehealth ASD assessments. There was an overall agreement between in-person and telehealth initial diagnostic impressions (66.7%). Accuracy varied with participant characteristics, as the telehealth assessments were more accurate in identifying true ASD diagnoses for younger and less verbal participants and more accurate in rejecting ASD diagnoses for older and more verbal participants. Regarding clinician confidence for diagnostic impressions, only a caregiver-reported ASD questionnaire significantly differed between certain and uncertain confidence levels of telehealth diagnostic impressions. These results indicate that more obvious ASD-related behaviors were related to higher clinician confidence in their telehealth diagnostic impressions. As for social validity responses, caregivers reported overall positive experiences of their telehealth experience. Family and participant characteristics did not impact the family's assessment experience. These preliminary findings support the continued use of telehealth ASD appointments as a valuable option for families seeking ASD evaluations. Telehealth assessments may facilitate timely diagnoses for all families, regardless of location and background. Future research is necessary to validate and expand on these preliminary findings.

Table of Contents

List of Tables

Introduction

Early assessments for autism spectrum disorder (ASD) are essential in reducing the long-term effects of a delayed diagnosis (Cidav et al., 2017; Hyman et al., 2020; Koegel et al., 2014; Turner & Stone, 2007; Warren et al., 2011). The "wait and see" approach is less efficient and has detrimental developmental and social consequences for children ultimately diagnosed with ASD (Koegel et al., 2014). Being diagnosed with ASD at an early age is linked to social communication gains, reduced occurrence of concerning behaviors (e.g., tantrums, aggression), and fewer co-occurring behavioral or mental health diagnoses (Koegel et al., 2014; Landa, 2018). Early intervention helps caregivers learn effective strategies for responding to behaviors and communicating with their child, thereby reducing family stress and secondary behaviors associated with ASD (e.g., tantrums, challenging behaviors; Koegel et al., 2014). Early intervention is most effective when caregivers and clinicians are heavily involved in intense and frequent intervention sessions at an early age (Landa, 2018). Ultimately, an early diagnosis of ASD decreases financial stress for the family and the community (Horlin et al., 2014; Koegel et al., 2014).

Barriers, such as difficulty accessing trained and knowledgeable clinicians, families' lack of understanding of the medical system, and possible mistrust of the medical system could ultimately delay the initial ASD evaluation (Best et al., 2021; Bishop-Fitzpatrick & Kind, 2017; Mandell et al., 2007; Martinez et al., 2018; Rosenberg et al., 2011; St. Amant et al., 2018). These barriers were further exacerbated during the COVID-19 pandemic, altering how people interacted with one another and modifying

how ASD evaluations were conducted (Berger et al., 2022; Gupta et al., 2020). Consequently, the methods of assessments for ASD were modified to prevent delays in initial diagnoses and services. This dissertation aimed to add to the growing literature on the accessibility and feasibility of telehealth ASD evaluations by analyzing telehealth ASD diagnostic tools compared to the traditional in-person diagnostic assessments. The clinicians' diagnostic impressions after completing telehealth ASD evaluations were explored, along with factors that influenced their clinical impressions and diagnostic confidence. Caregivers' experiences were examined to assess the feasibility of both telehealth and in-person appointments.

Autism Spectrum Disorder

Autism spectrum disorder (ASD) is a lifelong heterogeneous neurodevelopmental disorder consisting of social communication deficits and the presence of restricted and repetitive behaviors (American Psychiatric Association [APA], 2013). Social communication challenges are consistent across different social contexts. Social communication difficulties include poor social skills, difficulties engaging in reciprocal conversations, utilizing and understanding verbal and nonverbal communication, and initiating and maintaining relationships. Restricted and repetitive behaviors (RRBs) include but are not limited to, stereotyped or repetitive movements with body parts or vocalizations, difficulty deviating from routines, strong interests, and sensory preferences. To be diagnosed with ASD, a person must show these symptoms at a young age (APA, 2013).

The prevalence of ASD has increased in recent years from 1 out of 54 people in 2016 to 1 in 36 children in 2018, as documented by the Centers for Disease Control and Prevention (CDC, 2020; Maenner et al., 2023). This increase in prevalence not only accentuates the importance of early identification to ensure timely appropriate services but also reflects the increasing priority placed on early identification. While ASD is pervasive across all racial and ethnic groups (APA, 2013; CDC, 2020), its prevalence is not uniform among different regions and racial and ethnic groups within the United States (Maenner et al., 2021; Mandell et al., 2009; Rosenberg et al., 2011).

A family's socioeconomic status, racial and ethnic background, caregiver's educational background, and the child's sex have been found to impact *when* a family was initially assessed for ASD and influenced the *outcome* of that assessment (Avlund et al., 2021; Bishop-Fitzpatrick & Kind, 2017; Daniels & Mandell, 2014; Durkin et al., 2010; Fountain et al., 2011; Hyman et al., 2020; Lockwood Estrin et al., 2021; Maenner et al., 2023; Mandell et al., 2009; Mazurek et al., 2014; Rosenberg et al., 2011). For instance, studies have shown that African American and Black individuals were more likely to receive an initial diagnosis later than White individuals (Maenner et al., 2021; Mandell et al., 2009). Moreover, Hispanic and Latinx individuals were less likely to be diagnosed with ASD compared to other racial and ethnic groups (Blacher et al., 2014; Maenner et al., 2021; Mandell et al., 2009). Service inequities were also more apparent for Hispanic and Latinx individuals (Blacher et al., 2019). Geographical location also significantly affects the age of initial ASD diagnoses. Certain regions and states provided more access to early evaluations and services, enabling earlier diagnoses for individuals

(e.g., California's regional centers assisted with a younger initial ASD diagnosis age compared to Missouri; Maenner et al., 2021). Furthermore, the timing of caregiver concerns about their child and the intensity of their child's symptoms influenced the initial ASD diagnoses (Daniels & Mandell, 2014; Mazurek et al., 2014). Children with more subtle cues or parents who had difficulty describing concerning behaviors to professionals experienced delays in receiving an initial ASD diagnosis. Other factors that postponed an initial ASD diagnosis were: limited access to trained professionals and different clinical impressions among professionals (Martinez et al., 2018).

ASD Assessment Process: Brief Overview

Children can be assessed for ASD as young as 18 months (Lord et al., 2012; Matson et al., 2008); however, the average age of an initial diagnosis is around four years old (Maenner et al., 2021). ASD diagnoses remain stable even when diagnosed as early as 19 months (Guthrie et al., 2013; Pierce et al., 2019). Standardized measures for ASD effectively identify children with more subtle ASD symptoms, even at a young age (Avlund et al., 2021). The American Psychiatric Association recommends that an ASD evaluation include collecting information from various sources (e.g., caregiver interviews), obtaining the child's developmental history, and a clinician observation, in conjunction with an ASD assessment (APA, 2013).

The gold standard assessment tool for ASD evaluations is the Autism Diagnostic Observation Schedule, Second Edition (ADOS-2). The ADOS-2 is a standardized assessment with predefined activities designed to elicit ASD behaviors (Lord et al., 2012). The ADOS-2 is comprised of different modules based on the child's language

abilities and chronological age. Module 1 is for those with limited speech (e.g., one-word vocalizations) and are older than 31 months. Module 2 is for those who use phrase speech, such as two to three-word phrases. Module 3 is administered to individuals with fluent speech (i.e., use “and” and “but” in sentences). Module 4 is used for individuals with fluent speech and older than 15 years old. The toddler module is administered to ambulatory children between 12 to 30 months old (Lord et al., 2012). The ADOS-2 is highly accurate in correctly identifying those with ASD, with sensitivity of the ADOS-2 across Module 1, 2, and 3 ranging from 0.82 to 0.90. The specificity of the ADOS-2 for Modules 1, 2, and 3 ranges from 0.62 to 0.90 (Dorlack et al., 2018).

After completing the ADOS-2, assessors code the observed behaviors and transfer some behavioral codes to the ADOS-2 algorithm. The ADOS-2 algorithm is different for each Module, and for some Modules, it is further stratified based on the child’s language production and chronological age. The scores calculated from this algorithm yield two classifications. One is the ADOS-2 Classification, which determines if the individual’s total score meets the cut-off score for non-spectrum, autism, or autism-spectrum (Lord et al., 2012). The second score is the ADOS-2 Comparison Score, which indicates the severity of ASD symptoms corresponding to that person's language and chronological age (Gotham et al., 2012; Lord et al., 2012).

Similarly, the gold standard caregiver interview is a semi-structured interview called the Autism Diagnostic Interview-Revised (ADI-R). The ADI-R consists of standardized interview questions, with the option to ask additional questions for more details when necessary. This interview assists clinicians in differentiating ASD from

other diagnoses with overlapping or similar symptoms (Lord et al., 1994). Trained experts administer these gold-standard assessments to identify ASD (Martinez et al., 2018). However, one limitation of administering these assessments is the time required for families and clinicians, as they can take forty minutes to an hour and a half each (Lord et al., 1994; Lord et al., 2012).

Barriers to ASD Assessments

Unfortunately, accessing standardized ASD evaluations in a timely manner is not always possible (Gordon-Lipkin et al., 2016; Malik-Soni et al., 2022; Stahmer et al., 2019; Zuckerman et al., 2017). The literature has emphasized that rural communities often face challenges, including a lack of clinical services, trained individuals to administer screening measures, inconsistency in the types of ASD screeners being administered, and an overall lack of education about childhood diagnoses (Antezana et al., 2017; Martinez et al., 2018; Rosenberg et al., 2011). Additionally, travel time to clinics posed another barrier to receiving timely ASD-related services (Martinez et al., 2018; Stahmer et al., 2019).

Other barriers to obtaining a formal autism evaluation, not specific to rural communities, included the cost of appointments, especially when a family did not have access to health insurance (Lewis, 2017). Families also reported a mistrust of medical providers, either due to previously missed ASD diagnosis, needing to see multiple professionals before learning of their child's diagnosis, or feeling shamed, blamed, or invalidated by clinical providers. These barriers were further impacted by the scarcity of professionals specializing in ASD who could address the family's concerns and long

waitlists at clinics (Bivarchi et al., 2021; Elder et al., 2016; Gordon-Lipkin et al., 2016; Lewis, 2017; Liptak et al., 2008; Malik-Soni et al., 2021; Martinez et al., 2018; Stahmer et al., 2019; Zuckerman et al., 2017). Similarly, families perceived the healthcare system as confusing and unclear, leading to appointment delays (Avlund et al., 2021; Bivarchi et al., 2021; Elder et al., 2016; Lewis, 2017; Martinez et al., 2018). Lastly, some families reported not seeing the benefits of receiving a formal medical diagnosis (Lewis, 2017) or lacking education about ASD (Zuckerman et al., 2017).

Family characteristics also impacted the age of the initial diagnosis. Children from lower socioeconomic households were less likely to receive a timely ASD diagnosis (Bishop-Fitzpatrick & Kind, 2017; Bivarchi et al., 2021; Durkin et al., 2010; Hyman et al., 2020; Mazurek et al., 2014; Rosenberg et al., 2011). Racially and ethnically minoritized families faced unique challenges. One such barrier was related to language, as English-speaking families received more ASD-related services than non-English-speaking caregivers (St. Amant et al., 2018). In addition, caregivers reported that services were not always provided in their primary or preferred language (Stahmer et al., 2019). The racial identity of families contributed to delayed initial diagnosis (Bishop-Fitzpatrick & Kind, 2017; Mandell et al., 2007; Rosenberg et al., 2011). For example, African American families were half as likely to receive an ASD diagnosis for their child at their first specialized clinical appointment than White families (Mandell et al., 2007). A similar result was replicated with Latinx families, who were less likely to have therapy hours completed or receive ASD-related services compared to non-Latinx families (Zuckerman et al., 2017), despite Latinx children displaying ASD related behaviors equal

to White children drawn from the same relative community (Blacher et al., 2014; Liptak et al., 2008).

Overall, barriers to high-quality assessments for families could delay initial diagnoses (Stahmer et al., 2019), ultimately delaying service receipt. However, children who received earlier services exhibited better prognoses later in life, including improved cognitive and language abilities (Hyman et al., 2020; Turner & Stone, 2007; Warren et al., 2011).

The Impact of COVID-19 on ASD Assessment and Screening

When the COVID-19 pandemic started, lockdowns were initiated, and social distancing was a mandated rule in most of the United States. Regulations were published to prevent the spread of COVID-19, such as wearing masks, and public locations that brought large groups of people together were closed (Gupta et al., 2020). The pandemic impacted how autism evaluations and services were conducted, regardless of the level of impairment (Isensee et al., 2022). For example, some clinics were closed entirely, whereas others mandated personal protective equipment (PPE) when testing clients. However, masks and other facial coverings deviated from the standardization procedures of most ASD evaluation assessments (Berger et al., 2022).

Thus, not only did ASD evaluations change, but discrepancies in access to general medical care were inevitable. During the pandemic, historically underserved groups were disproportionately affected by either lack of medical care or distrust of healthcare providers (Best et al., 2021; Tai et al., 2020). Consequently, there were higher death rates and reduced contact with medical care for historically underserved groups (Tai et al.,

2020). As for families with a child with ASD, services either ceased or became sporadic. Only a small portion of services offered alternative support to families through telephone or telehealth means. Families who agreed to virtual support reported benefits from these services (Eshraghi et al., 2020; Isensee et al., 2022; Neece et al., 2020). According to caregiver reports, caregivers of children with developmental disorders, including ASD, experienced higher stress levels than other caregivers during the COVID-19 pandemic (Eshraghi et al., 2020; Isensee et al., 2022; Neece et al., 2020).

ASD Telehealth Assessments

Modified services were a potential solution to provide equitable services. Telehealth appointments offer services to a wide range of people in different areas from various professionals (Alfuraydan et al., 2020). These services were provided at a lower cost to families and reduced the time to receive an initial diagnosis through shorter waitlists (Alfuraydan et al., 2020; Gibbs et al., 2021; Smith et al., 2017; Stainbrook et al., 2019). This novel assessment style allowed families from diverse socioeconomic, racial and ethnic backgrounds equitable access to ASD-related services (Berger et al., 2022; Jang et al., 2022; Ludwig et al., 2021; Micheletti et al., 2023). Telehealth assessments also helped mitigate the long-term consequences of the COVID-19 pandemic, such as reducing appointment waitlists and backlogs (Ameis et al., 2020).

Different approaches were developed to assess children with ASD through telehealth. Common elements of ASD telehealth appointments included specific behaviors to observe, activities included in the assessments, and caregiver preparation for

the appointment (Berger et al., 2022; Stavropoulos et al., 2022). A few examples of telehealth ASD assessments are described below.

Video Recording

The video recording method involves storing videos of an individual and then sharing the videos with an expert clinician (Smith et al., 2017). One such assessment system is the Naturalistic Observation Diagnostic Assessment (NODA). As part of the NODA assessment, developmental history was collected through caregiver questionnaires, caregiver interviews were conducted, and video recordings of the individual were collected. Each video recording was taken for ten minutes in four specific scenarios. Three of the four scenarios were recorded at specific times (i.e., mealtime, playtime with someone, playtime alone). The fourth scenario allowed caregivers to record a time when their child displayed concerning behaviors. Caregivers received 30 minutes of training on the assessment tool before using the application. After the recordings were submitted, coders noted atypical behaviors in the videos. These tagged behaviors were then compared to the ASD criteria in the DSM 5. The average age of children administered the NODA were around four years old (Smith et al., 2017).

Although specificity and sensitivity were high when comparing diagnostic impressions from NODA to in-person appointments, NODA was not proposed as a replacement for the in-person ASD assessment procedures (Nazneen et al., 2015; Smith et al., 2017). Instead, NODA is a tool that provides detailed information about a child for future evaluations. This assessment tool is proposed to help inform other professionals about specific behaviors related to ASD (Nazneen et al., 2015; Smith et al., 2017). The

sensitivity of this measure was measured to be 0.85, with a 95% confidence interval of 0.67 and 0.94, and specificity was 0.94, with a 95% confidence interval of 0.73 and 0.92 (Smith et al., 2017).

Video Conferencing

Videoconferencing allows trained professionals to observe the client in real-time through video platforms. With the assistance of caregivers, clinicians provide instructions or prompts to caregivers to facilitate specific interaction scenarios (Alfuraydan et al., 2020). Caregivers are either trained before the appointment or provided instructions during the evaluation. Below are the three most common video conferencing telehealth evaluation tools.

Brief Observation of Symptoms of Autism (BOSA; Lord et al., 2020)

The BOSA is a modified version of an ASD diagnostic assessment, the ADOS-2 (Lord et al., 2012) and the Brief Observation of Social Communication Change (BOSCC; Grzadzinksi et al., 2016) that can be administered at a clinic or through telehealth. The BOSA was intended to emulate the ADOS-2; thus, it can be administered to individuals older than 15 months old with varying language abilities. The BOSA administration takes approximately 12 to 14 minutes and consists of predetermined activities in a set order with general guidelines associated with each activity. The materials for this assessment are standardized for each activity. Therefore, the materials for this assessment were sent to the family before the virtual appointment. Noteworthy behaviors related to ASD parallels those observed in the ADOS-2, such as social skills (e.g., quality of conversation, social initiations), nonverbal communication (e.g., gestures, eye contact),

and RRBs (Dow et al., 2021; Lord et al., 2020). Lord and colleagues (2020) investigated the sensitivity and specificity of the BOSA when using cut-off scores, similar to the ADOS-2 protocol algorithm. Across all Modules, including the Toddler Module, the sensitivity ranged from 86% to 100%, while the specificity ranged from 70% to 100%. The BOSA was less accurate at correctly identifying individuals with more language abilities than the ADOS-2 (Dow et al., 2021; Lord et al., 2020).

The Observation of Play Screener – Home Edition (OOPS: HE; Nickel, 2020)

The Observation of Play Screener – Home Edition (OOPS: HE) is another telehealth assessment used to assess ASD (Nickel, 2020). The administration takes approximately 25 to 30 minutes and includes ten different activities. It is administered to children between 12 to 36 months old. During live administration, the clinician provides prompts and directions to the caregiver. Additionally, caregivers were provided a list of materials and information to prepare before the appointment. Scores were calculated from this assessment and combined with a caregiver interview. Additionally, home videos were reviewed to obtain more information about the child (Berger et al., 2022; Nickel, 2020). Sensitivity and specificity were not reported for this measure.

A Telemedicine-Based ASD Evaluation Tool for Toddlers and Young Children (TELE-ASD-PEDS; Wagner et al., 2021b)

TELE-ASD-PEDS is administered virtually for approximately 10 to 20 minutes to assess children under 36 months with limited language. The administration consists of eight activities led by caregivers, with live coaching provided by the observing clinician. Caregivers also assist with describing and clarifying child behavior (Wagner et al.,

2021b). Before the evaluation, caregivers are instructed on the materials needed for the appointment. The materials for the evaluation are modified based on what the family has access to at their home. After the assessment, clinicians score the child's behaviors using dichotomous and Likert scores (Corona et al., 2020; Wagner et al., 2021b). The scores helped assessors determine the risk of ASD. A score of 11 or higher is considered "at risk" for ASD (Wagner et al., 2021b). The sensitivity and specificity of TELE-ASD-PEDS were calculated to be 0.80 and 1.00, respectively (Jones et al., 2022).

Limitations

Although virtual assessments allowed clinicians to see a wide range of clients regardless of their geographic location, there were limitations to this assessment method. For example, families of lower socioeconomic status generally had less access to electronic devices and the Internet to participate in these evaluations (Ameis et al., 2020; Baweja et al., 2022; Kryszak et al., 2022; Wagner et al., 2021a). If families had access to the technology, telehealth appointments required caregivers to understand the technology and to be in a location with a stable WIFI connection. Additionally, caregivers had to be mindful of the camera angle for the clinician to observe the child and their respective behaviors (Corona et al., 2021; Gibbs et al., 2021; Ludwig et al., 2021; Wagner et al., 2021a).

Due to the brevity of telehealth assessments, additional information was often needed outside of the ASD assessment. In conjunction with a telehealth ASD evaluation, caregiver interviews, and outside reports were typically obtained (Alfuraydan et al., 2020; Dow et al., 2021; Nazneen et al., 2015; Smith et al., 2017). Additional information was

needed, especially when the child presented with more severe ASD behaviors or had lower cognitive abilities. Deficits in attention and social motivation also impeded a clinician's diagnostic impressions, as it was difficult to eliminate distractions in the home environment or redirect a person's fleeting attention (Ludwig et al., 2021).

Assessment administration was altered compared to in-person administration since clinicians relied on caregivers to provide social bids and perform specific activities with the child. This strong reliance on caregivers required caregiver training before the appointment or live coaching during the appointment (Ludwig et al., 2021). Caregivers were also responsible for responding to challenging behaviors without the help of other professionals (e.g., tantrums; Ludwig et al., 2021). In addition to administration, the scoring of the assessment was different. Clinicians typically used their expertise to develop a final diagnosis rather than rely on a cut-off score that indicated the extent to which observed behaviors were associated with ASD, as provided by the ADOS-2 (Corona et al., 2020; Ludwig et al., 2021; Smith et al., 2017; Wagner et al., 2021b).

Benefits

Despite these limitations, there were several benefits of telehealth assessments. Firstly, clinicians saw a wide range of families across socioeconomic, geographic, and cultural backgrounds in their homes (Alfuraydan et al., 2020; Ameis et al., 2020; Antezana et al., 2017; Gibbs et al., 2021; Kryszak et al., 2022; Ludwig et al., 2021; Wagner et al., 2021a). The feasibility of this assessment allowed the evaluation process to be easily accessible to a wide range of families. Additionally, it allowed families to receive an assessment from various professionals of different backgrounds (Alfuraydan et

al., 2020; Ameis et al., 2020; Baweja et al., 2022; Gibbs et al., 2021). More specifically, this permitted children with more impairing ASD behaviors to be assessed sooner (Alfuraydan et al., 2020).

Telehealth allowed clinicians to observe the child at home and in different home scenarios, such as eating in the kitchen or playing in the living room. This flexibility also provided space for families to address their concerns and take a more pivotal role in the assessment process (Ameis et al., 2020; Corona et al., 2021; Kryszak et al., 2022; Ludwig et al., 2021). The transient nature of the telehealth assessments also allowed clinicians to see multiple families quicker, reducing the number of families placed on waitlists for evaluations (Alfuraydan et al., 2020; Ameis et al., 2020; Corona et al., 2021; Smith et al., 2017). A shorter assessment also resulted in cheaper assessment options for ASD evaluations (Dow et al., 2021). For example, families saved money by not taking time off from work, not commuting to a clinic, and not needing to hire childcare for other children in the home (Kryszak et al., 2022; Sutherland et al., 2018; Wagner et al., 2021a). Reese and colleagues (2015) reported that families saved approximately $35 in clinic travel costs.

Most importantly, caregivers perceived this process positively and were satisfied with the telehealth process (Corona et al., 2021; Reese et al., 2013; Sutherland et al., 2018). Some caregivers reported that their child would have been too shy to interact with new people at the clinic, thus providing the clinician with more information and behavioral observations than might have been obtained at an in-person evaluation (Corona et al., 2021). Clinicians also generally described high confidence in their final

diagnostic impressions of their clients (Gibbs et al., 2021; Wagner et al., 2021a; Wagner et al., 2022).

Although telehealth assessments are new, the research supporting telehealth evaluations is growing. Research has demonstrated that telehealth was just as accurate as in-person assessments. Regarding clinical judgment, clinicians generally agreed with diagnostic impressions when comparing in-person to telehealth appointments (Corona et al., 2023; Reese et al., 2013).

Initial Views of Telehealth: Screening and Diagnosis for Autism

Caregiver Perception

Caregivers described equivalent satisfaction with telehealth compared to in-person appointments. Regarding the instructions for telehealth appointments, caregivers reported that the instructions provided during the live telehealth appointment were understandable (Reese et al., 2013). According to caregiver reports, the telehealth assessment was a comfortable experience for families (Corona et al., 2021; Gibbs et al., 2021). Families perceived the telehealth experience as professional and were able to develop rapport with the assessment teams. Caregivers reported the clinicians as knowledgeable and effective communicators (Pompa-Craven et al., 2022; Talbott et al., 2022). Telehealth appointments were also convenient. Telehealth allowed families to avoid commuting to clinics while also being evaluated in a comfortable space that helped caregivers feel more relaxed during the assessment process (Gibbs et al., 2021; Talbott et al., 2022).

However, some caregivers reported that the technology may have hindered their experience. For example, some families reported that the screen was too small for the family to see the clinician or there were audio difficulties. In addition to technical challenges, caregivers thought video conferencing limited what the clinician observed during the appointment or that more subtle behaviors went unnoticed (Corona et al., 2021; Gibbs et al., 2021; Jones et al., 2022; Talbott et al., 2022). Although the reports on the telehealth appointments were generally positive, caregivers suggested that telehealth appointments should be offered as the first appointment, with the option to see a clinician in-person as the second appointment (Corona et al., 2021). Despite these limitations, most caregivers reported that they would participate in telehealth appointments again (Gibbs et al., 2021). However, when directly asked, caregivers (seven out of 11) reported a preference for in-person appointments (Jones et al., 2022).

Clinician Perception

Clinicians had mixed opinions on telehealth ASD evaluations. Most clinicians reported they were confident with their final diagnosis from the telehealth appointment (Gibbs et al., 2021; Wagner et al., 2021a; Wagner et al., 2022). However, this confidence level depended on individual factors, such as age, level of impairment, and language abilities (Gibbs et al., 2021; Kryszak et al., 2022; Wagner et al., 2021a; Wagner et al., 2022). Some clinicians reported that the younger the individual was, the less confident they were with their clinical impression (Gibbs et al., 2021). However, other clinicians reported higher diagnostic confidence with younger individuals with cognitive impairments and less diagnostic confidence with older individuals with more subtle ASD

behaviors (Kryszak et al., 2022; Wagner et al., 2021a; Wagner et al., 2022). Other behaviors that impeded clinical confidence were nonverbal behaviors that were hard to capture through telehealth, like eye contact (Gibbs et al., 2021) and overly active or anxious behaviors (Kryszak et al., 2022). On the other hand, some behaviors, such as social skills, were easier to assess and observe through telehealth (Gibbs et al., 2021).

Similar to caregiver experiences, clinicians also raised technology concerns. A frequent technology difficulty reported by clinicians was when families used a small device, like a phone, as it made it challenging to observe the individual's behaviors (Gibbs et al., 2021). It was also reported that families of lower socioeconomic status had more difficulty accessing the Internet or WIFI and lacked an understanding of how to use technology (Kryszak et al., 2022). Despite these limitations, clinicians reported that telehealth allowed them to gain valuable information about the client in a safe and comfortable space (Kryszak et al., 2022; Wagner et al., 2021a; Wagner et al., 2022). However, if the child demonstrated interfering and challenging behaviors (e.g., tantrums), the clinician relied on the caregiver to assist (Kryszak et al., 2022). Lastly, clinicians reported that caregivers sometimes interpreted directions and instructions differently, thus altering some of the assessment activities (Wagner et al., 2021a).

Overall, clinicians described positive feedback to using telehealth for interviews and feedback appointments. Moreover, telehealth led to fewer canceled appointments (Frank et al., 2021; Kryszak et al., 2022; Wagner et al., 2021a; Wagner et al., 2022). Most clinicians obtained enough information and behavioral observations from the telehealth assessment to provide an appropriate diagnosis to the client. It was noted that a

child who was difficult to diagnose through telehealth was also complicated to diagnose when seen in-person. Related to complex cases, the appointment was more multifaceted when a translator was present. However, this difficulty is also reported for in-person appointments (Kryszak et al., 2022). Due to the novelty of this assessment style, there is more to learn about the effectiveness of telehealth services. Therefore, it is important to understand for whom telehealth is most effective. Understanding which client's telehealth will be inadequate for obtaining a diagnostic impression is also critical.

Current Study Objective

Telehealth services must be tested in a community of diverse socioeconomic, cultural, and racial backgrounds, and with families who might not otherwise have equitable access to educational, psychological, or medical services (Frank et al., 2021; Khairat et al., 2019). The Inland Empire in Southern California consists of populations at high risk of receiving subpar health services compared to neighboring counties (Institute, 2020). Additionally, the Inland Empire is a medically underserved community, as the number of providers does not match the community's needs (Beckett & Morrison, 2010). As such, families in the Inland Empire lack access to high-quality health services and, therefore, often need to travel long distances to receive healthcare (Love et al., 2019).

The use of telehealth services increased during the COVID-19 pandemic, especially in historically underserved communities (Campos-Castillo & Anthony, 2021). Telehealth assessments provided an assessment tool that helps families with limited access to healthcare systems, who could not travel, or take time off from work to attend clinic appointments. With this in mind, it is necessary to understand telehealth ASD

evaluations for individuals of all ages and diverse backgrounds. Additionally, the ASD feasibility and acceptability of this novel assessment tool within rural and medically underserved communities. It is also important to understand which child characteristics are best suited for telehealth ASD evaluations.

Lastly, telehealth ASD evaluations have been frequently used with a younger population (e.g., younger than four years old) to determine ASD risk levels (Nickel, 2020; Smith et al., 2017; Wagner et al., 2021b). Limited ASD telehealth assessments have been used with older populations with more language abilities. Although the Brief Observation of Symptoms of Autism (BOSA) protocol is appropriate for older and more verbal children, this assessment requires standardized materials for administration, thus, limiting who can be assessed with this assessment tool (Lord et al., 2020). An ASD telehealth assessment tool with more flexible administration guidelines for older and more verbal individuals should be developed and evaluated.

Research Questions

The following questions were analyzed as part of this study:

1. What is the efficacy and reliability of diagnosing ASD using telehealth compared to traditional in-person evaluations?
    a. In what proportion of cases was there congruence across telehealth and in-person assessments for final diagnoses?
    b. What are the sensitivity and specificity of the telehealth protocols?
    c. Are there differences in clinician confidence levels between telehealth and in-person evaluations?

d. Were specific child characteristics related to clinician confidence levels at telehealth appointments?

2. Are there differences in caregiver-rated social validity responses based on the assessment procedure (i.e., telehealth vs. in-person)?
    a. What caregiver and child characteristics are associated with caregiver social validity ratings from telehealth appointments (e.g., SES, racial/ethnic background)?

Methods

This dissertation is part of a larger research grant that aims to understand the efficacy and reliability of telehealth ASD evaluations. The ASD telehealth evaluation tools that have been developed and described below have not been tested in a low income rural population and with historically marginalized communities. This dissertation is an analysis of the first 27 participants of that ongoing study.

Participants

Participants were recruited through the UCR SEARCH center. The UCR SEARCH Center is a free autism evaluation community clinic in the Inland Empire in Southern California. Families were referred to the SEARCH Center by other community professionals and clinics in the area or through online searching. Families with concerns regarding ASD contacted the SEARCH Center. An initial phone intake was conducted using the Social Communication Questionnaire (SCQ; Rutter et al., 2003). This measure was used to screen each participant to determine if the individual showed several ASD-like behaviors. The SCQ recommends a cut-off score of 15. However, the SCQ has

higher sensitivity and specificity when the SCQ cut-off score is lowered to 11 with young children (Wiggins et al., 2007). Families were typically placed on the waitlist if a child scored a ten or higher on the SCQ, however, this approach was not always followed rigidly. Other recommendations were provided to families if the child displayed other behaviors not related to ASD even if they scored high on the SCQ.

As previously stated, participants are still being recruited for this study, as such all participants who completed both assessment time points, both telehealth and in-person, by March 1, 2023, were considered for this study. Participants who partook in this study were families who were in the process of seeking an ASD evaluation. Therefore, no active recruitment occurred, as any family seeking an ASD evaluation at this clinic was provided this research opportunity. This study has been approved by the IRB at the University of California, Riverside.

As shown in Table 1, 27 participants were assessed and completed both appointments before March 1, 2023. Of the 27 participants, 81.5% were male and had an average age of 86.8 months (SD = 42.8 months). Most participants spoke English as their primary household language (96.3%) and 37.0% identified as being Hispanic and/or Latinx. Table 2 indicates child characteristics after randomization to a telehealth appointment or an in-person appointment first.

Table 1

Participant Characteristics

Child Characteristics (N = 27)	
Sex (% males)	81.5%
Age (months)	86.8 (42.8)
SCQ Total Score	22.3 (6.5)
Primary Household language	
English	96.3%
Spanish	3.7%
Race/Ethnicity	
Asian	3.7%
African American/Black	11.1%
Hispanic/Latinx	37.0%
White	29.6%
Mixed	14.8%
Other	3.7%

Table 2

Participant Characteristics by First Appointment

	Child Characteristics (N = 27)	
	Telehealth First Appointment (N = 11)	In-Person First Appointment (N = 16)
Sex (% males)	90.9%	75.0%
Age (months)	95.3 (50.1)	81.0 (37.6)
SCQ Total Score	22.3 (7.5)	22.2 (5.1)
Primary Household language	100.0%	93.8%
English	0.0%	6.3%
Spanish		
Race/Ethnicity		
Asian	9.1%	0.0%
African American/Black	18.2%	6.3%
Hispanic/Latinx	18.2%	50.0%
White	36.4%	25.0%
Mixed	9.1%	18.8%
Other	9.1%	0.0%

Table 3 provides the demographic information of the participant's caregivers. Most caregivers were the participant's biological parent (72.2%). In addition, 37.0% of the participant's families' incomes were less than $50,000, with most families having four family members, including the caregiver and the participant, dependent on the income (40.7%).

Table 3

Caregiver Characteristics

Caregiver Characteristics	Caregiver (N = 54)
Family Income (% <50,000)	37.0%
People Dependent on Family Income	
Two	7.4%
Three	33.3%
Four	40.7%
Five	14.8%
Other	3.7%
Age (years)	39.13 (9.25)
Relationship to Child	
Biological	72.2%
Adoptive	9.3%
Other	11.1%
Missing	7.4%
Race/Ethnicity	
Asian	3.7%
African American/Black	7.4%
American Indian/Alaska Native	1.9%
Hispanic/Latinx	35.2%
White	33.3%
Mixed	9.3%
Other	1.9%
Missing	7.4%
Highest Level of Education	
Less than High School	5.6%
High School/GED	25.9%
Associative Degree/Some College	33.3%
Bachelor's Degree	16.7%
Graduate Degree	9.3%
Missing	9.3%

Procedure

After completing the intake phone call and determining that the child received a score of 10 or higher on the SCQ (Rutter et al., 2003), families were invited to participate in the research study and informed about the study's objectives and procedures.

Consented caregivers were informed that their child would undergo two assessments – one in-person at the UCR SEARCH Center and the other online via Zoom, and that they would be randomly assigned to the first condition. The two appointments were typically scheduled three weeks apart. The time between in-person and telehealth appointments ranged from 11 to 78 days, with an average of 24.85 days (SD = 14.04 days). At each appointment, child participants were administered one cognitive assessment and an ASD evaluation. Further details about the telehealth and in-person appointments are described later in this section.

Following each assessment, families completed an online social validity questionnaire to understand their assessment experience. After both the telehealth and in-person appointments were completed, families filled out another online questionnaire to compare their experiences from each appointment. After completing both appointments, a written report describing the in-person assessment results was provided to families, and a feedback session was scheduled either by phone or a Zoom video conference, based on caregiver preference. At each appointment, an ADOS-2 research reliable clinician was present. All assessments provided at each appointment are shown in Table 4.

Table 4

Assessments Administered at Appointments

	Telehealth	In-Person
Cognitive	WPPSI-IV (for children <5 years) DAS-II (for children >5 years)	WPPSI-IV (children < 7 years) WASI-II (for children >7 years)
ASD	TELE-ASD-PEDS TELE-ASD-KIDS (Module 2 & 3)	ADOS-2
Caregiver Questionnaires	N/A	Vineland-3 Domain Level CBCL SRS-2 SSIS

Note. Only one cognitive assessment was administered at each appointment. The cognitive assessment administered was dependent on the participant's age as shown in the table. WWPSI-IV = Wechsler Preschool and Primary Scale of Intelligence – Fourth Edition; DAS-II = Differential Ability Scales – Second Edition; WASI-II = Wechsler Abbreviated Scale of Intelligences, Second Edition; TELE-ASD-PEDS = A Telemedicine-Based ASD Evaluation Tool for Toddlers and Young Children; ADOS-2 = Autism Diagnostic Observation Schedule, Second Edition; CBCL = Child Behavior Checklist; SRS-2 = Social Responsiveness Scale, Second Edition; SSIS = Social Skills Improvement System Rating Scale.

In-Person Appointment

At the in-person appointment, families commuted to the SEARCH clinic in Riverside. Children first underwent a full cognitive assessment, followed by the ADOS-2 (Lord et al., 2012). Based on the cognitive assessment's age cut-off, children under seven were administered the Wechsler Preschool and Primary Scale of Intelligence – Fourth Edition (WPPSI-IV; Wechsler, 2008), while those over seven years old were administered the Wechsler Abbreviated Scale of Intelligence, Second Edition (WASI-II; Wechsler, 2011). All subtests necessary for a Full Scale IQ (FSIQ) score were administered. The ADOS-2 was administered or overseen by an ADOS-2 research reliable assessor (Lord et al., 2012). The assessment team scored the cognitive and

ADOS-2 before developing an initial diagnostic impression of the participant. The initial diagnostic impression did not include other caregiver questionnaires completed at the in-person appointment.

Caregivers completed additional questionnaires about participants' behaviors and social abilities at the in-person appointment. These included the Vineland-3 Domain Level, Child Behavior Checklist (CBCL), Social Responsiveness Scale, Second Edition (SRS-2), and Social Skills Improvement System Rating Scale (SSIS). At the end of the in-person appointment, caregivers completed a social validity questionnaire. These caregiver questionnaires were scored after the participants completed the in-person appointment.

Telehealth Appointment

For the telehealth appointment, caregivers were emailed a list of materials for their telehealth appointment and provided Zoom instructions before the appointment. At the telehealth appointment, the cognitive assessment was conducted before the ASD evaluation to emulate an in-person appointment structure. Depending on the age of the participant, the Differential Ability Scales, 2nd Edition (DAS-II; Elliott, 2007) or the Wechsler Preschool and Primary Scale of Intelligence – Fourth Edition (WPPSI-IV; Wechsler, 2008) were administered virtually. If the child was between two years six months and three years 11 months, two subtests from the WPPSI-IV were administered (Receptive Vocabulary and Information) to obtain a Verbal Comprehension Composite score. Children between the ages of four and five years old were administered four subtests from the WPPSI-IV (Similarities, Picture Concepts, Information, and Matrix

Reasoning). Individuals older than five years old were administered four subtests from the DAS-II to obtain a Verbal Ability Cluster score (Verbal Similarities and Word Definitions) and the Nonverbal Ability Cluster score (Matrices and Sequential and Quantitative Reasoning). These tests were chosen by their ability to be administered virtually (e.g., no manipulatives, pictures that can be shared through video conferencing). The DAS-II was administered to present a novel cognitive assessment at the telehealth appointment when possible.

The ASD assessment chosen for the telehealth appointment was determined by the participant's language level, similar to the ADOS-2. Children with no language or one-word phrases were administered the TELE-ASD-PEDS (Wagner et al., 2021b). Participants with phrase speech (e.g., two or three-word phrases) or fluent speech (e.g., used "and" or "but" in sentences) were administered an adapted version of the TELE-ASD-PEDS, called TELE-ASD-KIDS (Stavropoulos et al., 2021). After completing the assessments, the telehealth assessment team scored the assessments and developed an initial diagnostic impression.

After completing the telehealth assessments, caregivers were asked if the observed behaviors accurately represented their child. Any additional concerns or questions were addressed, and a social validity questionnaire was provided to understand the family's telehealth experience via email.

Diagnostic Impressions

After the ASD evaluation was completed, whether in-person or via telehealth, clinicians were required to answer several questions regarding their initial diagnostic

impression of the participant. The first question prompted clinicians to consider whether they would diagnose the participant with ASD. The available options were "Yes," "Unsure," and "No." Next, clinicians were asked to indicate their confidence level in their initial diagnostic impression using the following scale: 1 = Completely Uncertain, 2 = Somewhat Uncertain, 3 = Somewhat Certain, and 4 = Completely Certain.

Furthermore, clinicians were asked to consider whether participants displayed other behaviors that influenced their assessment results, such as anxiety, overactivity, or tantrums. This question aimed to determine if other disruptive behaviors might have contributed to elevated scores in the ASD evaluation when unrelated ASD behaviors and symptoms were present. Three response options were provided for this question: "Other behaviors common in autism might have biased a few codes (e.g., overactive/inattention)," and "Other behaviors (not ASD) heavily bias the codes above, and they do not reflect the clinical impression of ASD (e.g., anxiety, attention difficulties, impulsive behaviors)."

In addition to these questions, clinicians were required to document their initial diagnostic impression by listing applicable DSM-5 diagnoses. Each assessment team formulated an initial diagnostic impression after their respective appointments. Both assessment teams received identical background information, including the phone intake, SCQ results, and their respective cognitive and ASD assessment results from their attended appointment.

Social Validity Questionnaires

After each appointment, caregivers completed an online social validity questionnaire, which assessed the effectiveness and feasibility of the recent assessment process. Examples of questions included: "This assessment process seemed easy" and "I felt that the assessor was able to collect important information about my child." After families completed both the telehealth and in-person appointments, caregivers were asked questions comparing their experiences from each appointment. The specific questions for each social validity questionnaire can be found in Table 11, Table 12, and Table 13.

Participants with more verbal abilities (i.e., those administered Module 3) were asked about their assessment experiences. After completing both appointments, the clinician asked the participants social validity questions and provided a visual aide to assist with their responses. The clinician recorded the participants' responses. The social validity questions presented to the participants can be found in Table 11.

Final Diagnosis

The final diagnosis relied on the initial diagnostic impression concluded at the in-person appointment, as this is the gold standard method of providing an ASD diagnosis. Like any clinic, complex cases occurred but were rare. For these complicated cases, both assessment teams met to review discrepancies between the in-person and telehealth diagnostic impressions, and reviewed other relevant information (e.g., previous reports, developmental history). The assessment scores from both appointments and caregiver measures were studied to determine the best-fitting final diagnosis.

Measures

ADOS-2

For the in-person ASD evaluation, children were administered the ADOS-2 as their ASD assessment (Lord et al., 2012). The ADOS-2 consists of five modules, Module 1, 2, 3, 4, and Toddler Module. The specific module administered was determined by the child's language abilities and chronological age, as outlined in the ADOS-2 manual. Individuals with limited language or one-word speech were administered a Module 1, which primarily involved play-based activities. For those with phrase speech, a Module 2 was administered, which included play activities and semi-structured conversations. Module 3 was administered to individuals with fluent speech (e.g., use "and" and "because" in sentences) and included more conversational activities with limited play-based activities.

The behavioral observations and codes were completed as instructed by the manual. The ADOS-2 algorithm was utilized to calculate the Social Affect, Restricted and Repetitive Behavior (RRB), Overall Total score (the sum of the Social Affect and RRB scores), ADOS-2 Calibrated Severity Score (CSS) and ADOS-2 Classification. The ADOS-2 Classification provides diagnostic classification of the individual resulting with either autism spectrum disorder, autism, or non-spectrum. The CSS indicates symptom severity of the individual that can be compared across modules. The CSS ranges from one to ten, with ten indicating more behaviors related to ASD. The CSS is determined by the individual's chronological age and Overall Total Score (Lord et al., 2012).

TELE-ASD-PEDS

TELE-ASD-PEDS (Wagner et al., 2021b) was developed to assess children younger than three years old with limited speech via telehealth. For our clinic, this protocol was administered to children of any chronological age with limited language. TELE-ASD-PEDS consists of ten activities and seven codes, including calling their child's name, directing their child's attention, and playing toys with their child. A complete list of the activities and codes of the TELE-ASD-PEDS can be found in Table 5 (Wagner et al., 2021a; Wagner et al., 2021b). Caregivers were actively involved in the administration of this assessment and received live coaching from the clinician during the activities. After the assessment, the clinician coded seven items using dichotomous (i.e., 1 and 3) and Likert scales (i.e., 1, 2, and 3), with higher numbers indicating more behaviors related to ASD. For dichotomous scores, a 1 implies no observed ASD-related behaviors, whereas a 3 indicates that ASD behaviors were observed. Similarly, for Likert scoring, a 1 meant no observed ASD-related behaviors, a score of 2 indicated ASD behaviors were observed but not at a clinical level, while a score of 3 signified behaviors consistent with ASD diagnostic criteria. Total scores were independently calculated for each dichotomous and Likert scale (Wagner et al., 2021a; Wagner et al., 2021b).

TELE-ASD-KIDS

TELE-ASD-KIDS (Stavropoulos et al., 2021) was adapted from elements from the TELE-ASD-PEDS and ADOS-2 protocol to be administered to children with more language abilities (Corona et al., 2020; Lord et al., 2012). This protocol includes a Module 2 and Module 3. The module was determined by the participant's language

abilities. Module 2 was administered to those with phrase speech, while those with fluent speech (e.g., “and” and “but” in sentences) were administered Module 3. The TELE-ASD-KIDS Module 2 comprised of 13 activities and nine codes, while TELE-ASD-KIDS Module 3 included nine activities and ten codes. Table 5 provides a list of activities and behavioral codes for each module. During the administration of TELE-ASD-KIDS Module 2, caregivers were present to assist with challenging behaviors and engage with the child in predefined activities (e.g., calling the child’s name, directing the child’s attention to objects in the room). Live coaching was provided by the clinician to the caregiver. For example, during Module 2, caregivers were coached on how to engage in a routinized play activity with their child (e.g., rolling a ball back and forth). In contrast, caregiver involvement was optional during the administration of Module 3.

The clinician developed behavioral codes based on their observations. Each behavioral code was scored using a Likert (i.e., one, two, or three) and dichotomous code (i.e., one or three). Higher scores for the behavioral code indicated more ASD-related behaviors. The codes were recorded similarly to the TELE-ASD-PEDS. The dichotomous scores indicated whether or not the ASD-related behaviors were observed. A score of 3 signified ASD-related behaviors were observed while a score of 1 meant ASD-related behaviors were not observed. For the Likert scoring, a 1 meant no behaviors observed were associated with ASD. A Likert score of 2 indicated ASD behaviors were observed but not at a clinical level, while a 3 code indicated behaviors were observed consistent with ASD diagnostic criteria.

Table 5

Telehealth ASD Assessment Activities and Behavioral Codes

	TELE-ASD-PEDS (No language or Single Words)	TELE-ASD-KIDS Module 2 (Phrase Speech)	TELE-ASD-KIDS Module 3 (Fluent Speech)
A	**Activity 1:** Cause & Effect **Activity 2:** Calling Name **Activity 3:** Direct Child' Attention **Activity 4:** Joint Play **Activity 5:** Calling Name **Activity 6:** Direct Child's Attention **Activity 7:** Familiar Play Routine **Activity 8:** Ready, Set, Go **Activity 9:** Snack **Activity 10:** Free Play + Ignore	**Activity 1:** Cause & Effect **Activity 2:** Calling Name **Activity 3:** Direct Child' Attention **Activity 4:** Joint Play **Activity 5:** Calling Name **Activity 6:** Direct Child's Attention **Activity 7:** Description of a Picture **Activity 8:** Conversation **Activity 9:** Demonstration Task **Activity 10:** Familiar Play Routine **Activity 11:** Ready, Set, Go **Activity 12:** Snack **Activity 13:** Free Play + Ignore	**Activity 1:** Description of a Picture **Activity 2:** Conversation **Activity 3:** Demonstration Task **Activity 4:** Cartoons **Activity 5:** Emotions **Activity 6:** Social Difficulties/Annoyance **Activity 7:** Friends, Relationships, Marriage **Activity 8:** Loneliness **Activity 9:** Light Conversation
C	• Socially Directed Speech • Frequent/ Flexible Eye Contact • Unusual Vocalization • Unusual Play • Unusual/ Repetitive Body Movement • Combines Gestures, Eye Contact, and Speech • Sensory Exploration/ Reaction	• Socially Directed Speech • Frequent/ Flexible Eye Contact • Unusual Vocalization • Unusual Play • Unusual/ Repetitive Body Movement • Combines Gestures, Eye Contact, and Speech • Sensory Exploration/ Reaction • Unusual Interests • Social Overtures	• Conversation • Volunteers Information • Frequent/ Flexible Eye Contact • Stereotypes Speech/ Vocalization • Facial Expressions • Combines Gestures, Eye Contact, and Speech • Social Overtures • Unusual/ Repetitive Body Movement • Unusual Interests • Sensory Exploration/ Reaction

Note. A = Activities, C = Codes.

Cognitive Assessments

Differential Ability Scales – Second Edition (DAS-II). The DAS-II is a cognitive assessment used for individuals between two years six months and 17 years 11 months. However, individuals between two years six months and three years five months receive the Early Years Battery, which produces a General Conceptual Ability (GCA) standard score, Verbal Ability cluster standard score, and Nonverbal Ability cluster standard score. Individuals older than seven years old are administered School-Age Battery which produces a GCA standard score, Verbal Ability standard score, Nonverbal Ability standard score, and Spatial Ability standard score (Elliott, 2007). For this study, the DAS-II's School-Age Battery subtests were administered to produce a Verbal Ability Cluster standard score. This subtest was only administered at the telehealth appointment to individuals older than five years old. Individuals younger than five years old were administered a WPPSI-IV at their telehealth appointments, described below.

Wechsler Preschool and Primary Scale of Intelligence – Fourth Edition (WPPSI-IV). The WPPSI-IV is a cognitive assessment administered to individuals aged between two years six months and seven years seven months. This assessment provides a FSIQ and three composite scores, including Verbal Comprehension, Visual Spatial, and Working Memory (Wechsler, 2012). For this study, individuals younger than five years old were administered a WPPSI-IV at their telehealth appointments to obtain the Verbal Comprehension standard score. Individuals younger than six years old were administered the entire WPPSI-IV protocol at their in-person appointment to obtain all composite scores and the FSIQ. Individuals older than six years old were administered the WASI-II

at their in-person appointment. The FSIQ collected at the in-person appointment was used in analyses.

Wechsler Abbreviated Scale of Intelligence, Second Edition (WASI-II). The WASI-II is a cognitive assessment administered to individuals older than six years old. After completing the assessment, a FSIQ, Verbal Comprehension Index, Perceptual Reasoning Index, Verbal Intelligence Quotient, and Performance Intelligence Quotient are developed (Wechsler, 2011). For this study, the WASI-II was only administered at the in-person assessment with individuals older than six years old. The FSIQ was used when in analyses.

Caregiver Questionnaires

Social Communication Questionnaire (SCQ). The SCQ is a caregiver-reported assessment used to evaluate ASD-related behaviors in individuals over four years old. It consists of 40 items with two answer options: yes or no. A total score of 15 or higher indicates behaviors associated with ASD, but researchers have found that lowering the cut-off score to 11 continues to have high sensitivity and specificity, especially with younger individuals (Wiggins et al., 2007). Behaviors related to ASD are counted and summed to create a total score for the assessment. The SCQ assessed ASD-related behaviors during the intake process to determine if the individuals displayed high levels of ASD symptoms for an ASD evaluation (Rutter et al., 2003).

Vineland-3 Domain Level. The Vineland-3 Domain Level assesses the participant's adaptive skills compared to same-aged peers and can be used for individuals older than three years old. This assessment produces standard scores, including adaptive

behavior composite (ABC), communication, daily living skills, socialization, and motor skills (Sparrow et al., 2005). Caregivers completed this questionnaire at the in-person appointment. The ABC score was used when investigating participant characteristics.

Data Analysis Plan

Research Question 1: Efficacy and Reliability of Telehealth Assessments

The first research question addressed the following: (1) What is the efficacy and reliability of diagnosing ASD using telehealth compared to traditional in-person evaluations? (1a) In what proportion of cases was there congruence across telehealth and in-person assessments for final diagnoses? (1b) What are the sensitivity and specificity of the telehealth protocols? (1c) Are there differences in clinician confidence levels between telehealth and in-person evaluations? (1d) Were specific child characteristics related to clinician confidence levels at telehealth appointments?

To understand the efficacy and reliability of diagnosing ASD using telehealth assessments, the initial diagnostic impressions and the clinicians' confidence levels for their diagnostic impression were compared between telehealth and in-person appointments.

For this question, terms like "ASD-ASD" or "no ASD-ASD" are utilized. These notations indicate the diagnostic impression from each appointment, with the initial diagnostic impression from the telehealth appointment indicated first. For example, if a participant's diagnostic impression from their telehealth appointment was "no ASD," but at their in-person appointment, the diagnostic impression was "ASD," this will be symbolized as "no ASD-ASD" category (i.e., false negative).

The initial diagnostic impressions from each appointment were divided into four categories and were denoted as follows: when the assessment teams from both in-person and telehealth agreed on an initial ASD diagnostic impression = ASD-ASD, both agreed that an ASD diagnosis was **not** the diagnostic impression = no ASD-no ASD, telehealth concluded with an initial ASD diagnosis and in-person did not = ASD-no ASD, and telehealth resulted with no ASD for their initial diagnosis and in-person did = no ASD-ASD. A percentage was developed for the four groups.

Child characteristics were used to understand which participant characteristics were more likely to result in agreement between the two assessment teams (ASD-ASD and no ASD-no ASD) and no agreement between the two appointments (no ASD – ASD and ASD – no ASD). It was originally proposed to investigate child characteristics among the following groups: ASD agreement (ASD – ASD), no ASD agreement (no ASD – no ASD), and diagnostic disagreement (no ASD – ASD and ASD – no ASD). However, due to the small sample size in each diagnostic impression group, the ASD-ASD and no ASD - no ASD groups (or diagnostic agreement groups) were combined for this analysis. As such, independent sample t-tests were used to understand the relationship between child characteristics and the agreement and disagreement among the two assessment teams. Child characteristics included the participant's age, FSIQ, Total Behavioral score, Vineland ABC, and SCQ Total Score.

The sensitivity and specificity of telehealth ASD assessments were also calculated. Sensitivity is the "proportion of true positives" (Altman & Bland, 1994, p. 1552; Chu, 1999). A true ASD diagnosis relied on the diagnostic impressions concluded

at the in-person appointment. For this study, the sensitivity of the telehealth protocols was calculated by using the following formula: $Sensitivity = \frac{ASD-ASD}{(ASD-ASD)+(\ no\ ASD-ASD)}$. In other words, the number of true ASD diagnoses was divided by the sum of true ASD diagnoses and false negative ASD diagnoses from telehealth appointments (i.e., those not provided an ASD initial diagnostic impression at their telehealth appointment but received an ASD diagnostic impression at the in-person appointment; Baratloo et al., 2015). Specificity is the "proportion of true negatives" (Altman & Bland, 1994, p. 1552; Chu, 1999). Specificity was calculated using the following formula: $Specificity = \frac{no\ ASD-no\ ASD}{(no\ ASD-no\ ASD)+(ASD-no\ ASD)}$. Specificity was calculated by dividing the true no ASD diagnoses by the sum of true no ASD initial diagnostic impressions and the total false positives diagnoses of ASD at the telehealth appointment (i.e., cases provided an ASD at the telehealth appointment but were not given an ASD diagnostic impression at the in-person appointment; Baratloo et al., 2015). The sensitivity and specificity were calculated separately for TELE-ASD-PEDS and TELE-ASD-KIDS.

This question also examined the confidence levels of the assessors' initial diagnostic impression and explored child characteristics that influenced the clinicians' confidence levels for telehealth evaluations. Clinician confidence scores were converted into numbers: 1 = Completely Uncertain, 2 = Somewhat Uncertain, 3 = Somewhat Certain, and 4 = Completely Certain. This score conversion followed the procedure used by Wagner and colleagues (2021a).

The child characteristics associated with clinician confidence levels for telehealth appointments were investigated. Two groups were created: participants where the

clinicians were either *Completely Uncertain* or *Somewhat Uncertain* with their initial diagnostic impression and those whose clinicians were *Completely* and *Somewhat Certain* with the initial diagnostic impression at the telehealth appointment. Child characteristics were evaluated between these two groups, including the participant's chronological age in months at their first appointment, number of previous diagnoses, FSIQ, Total Likert score from their telehealth ASD assessment, Vineland ABC, SCQ total score, and Total Behavioral Score (e.g., anxiety, aggression, overactivity). Independent-sample t-tests and Chi-square test of homogeneity analyzed group differences. The child's assigned sex at birth was not included in the analysis due to the limited number of females included in the sample.

The Total Behavioral Score was calculated by combining the other behavior scores as reported by the assessment team from both appointments into one code to capture the child's overall disruptive behaviors across both settings. Clinicians rated other non-ASD behaviors that impacted their ASD evaluation after completing the assessment. Combining these scores yielded a score that ranged from two to six. A score of two meant no other behaviors influenced the assessment codes. A six meant other behaviors that were not related to ASD were present during the evaluation and impacted the codes scored for the autism evaluation (e.g., anxiety, tantrums, aggression, overactivity).

A Fisher exact test also analyzed differences in clinician confidence levels between telehealth and in-person appointments.

Research Question 2

The second research question examined the social validity questionnaires completed by caregivers. This research question aimed to answer the following questions (2) Are there differences in caregiver-rated social validity responses based on the assessment procedure (i.e., telehealth vs. in-person)? and (2a) What caregiver and child characteristics are associated with caregiver social validity ratings from telehealth appointments (e.g., SES, racial/ethnic background)?

A Paired-Sample t-test compared the mean scores between each questionnaire statement. The current Likert scale was converted to numbers; 1 = Strongly Disagree, 2 = Disagree, 3 = Undecided/Neutral, 4 = Agree, and 5 = Strongly Agree. Corona and colleagues (2021) used a similar analysis for their caregiver social validity statements.

Descriptive statistics were completed for caregivers' responses about their experiences at each appointment. Caregivers were also asked to compare their experiences from telehealth and in-person appointments. Percentages and descriptive statistics were calculated based on each statement provided to families.

Child participants with more language abilities who had fluent speech (e.g., sentences that included "and" or "but") or those who were administered a Module 3 were also asked about their assessment experience. Percentages and descriptive statistics of the participants' responses were calculated for each statement provided.

The second part of this research question explored participant and family characteristics associated with telehealth social validity ratings. Matthews and colleagues (2021) studied caregivers' telehealth experience as reported by social validity

questionnaires. The research team first averaged all social validity questions to produce an acceptability summary score. The acceptability summary score compared family responses to family and child characteristics. This question replicated the statistical analysis Matthews et al. (2021) completed. As such, the telehealth social validity questions were averaged to create a Total Social Validity score, this study's label for Matthews et al. (2021) acceptability summary score. Like Matthews and colleagues (2021), reverse coding was necessary for one item. The one question that was reversed coded was, "I needed to learn a lot of things about technology before I could start this telehealth assessment," as it was the only negatively phrased statement. The Total Social Validity score was correlated to participant and family characteristics. The analysis did not include participants whose caregivers did not complete the social validity questionnaires. Spearman's rank-order correlations were conducted to compare the relationship between the Total Social Validity score and the following continuous child and family characteristic variables: adaptive skills (Vineland-3 ABC), cognitive functioning (FSIQ from the WPPSI-IV or WASI-II), participant's Total Other Behavior score (total score of other behaviors observed during assessments), SCQ Total Score, telehealth ASD assessment Likert total score, telehealth clinician confidence score on initial diagnostic impressions, and the time traveled to the clinic.

A point-biserial correlation was conducted to compare the relationship between the Total Social Validity score and categorical variables, including family's income and racial and ethnic background (Kornbrort, 2014). The child's sex was not included in this analysis due to the limited number of females in the data. The family's income was

divided into two groups: $50,000 or less and greater than $50,000. The child's racial and ethnic background was separated into two groups: non-Hispanic/Latinx White origin and Hispanic/Latinx origin. Caregivers reported if their child was of Hispanic, Latino/a, or Spanish origin with one of the following answers: yes or no. The caregiver's response was used for this analysis. This analysis produced a correlation coefficient that described the strength of the relationship between the two variables being assessed (Schober et al., 2018; Sedgwick, 2012).

Results

Before conducting analyses, the effects of randomization were assessed. Participant characteristics were compared to determine whether significant differences existed between individuals randomized to telehealth (N = 11) and those randomized to an in-person appointment (N = 16) first. Out of the total participants, 16 out of 27 (59.3%) were initially assessed in-person. Independent sample t-tests and Chi-square test of homogeneity were used to compare the effects of randomization.

There were no significant differences between the two groups regarding the participant's age in months at the first appointment (telehealth M = 95.3, SD = 50.1; in-person M = 81.1, SD = 37.5; $t(25) = -0.8$, $p = 0.41$), sex assigned at birth ($X^2(1) = 1.1$, $p = 0.30$), participant's FSIQ (telehealth M = 79.2, SD = 17.7; in-person M = 93.8, SD = 25.5; $t(20) = 1.5$, $p = 0.14$), participant's race/ethnicity (non-Hispanic background compared to Hispanic background; $X^2(1) = 0.4$, $p = 0.53$), SCQ total score (telehealth M = 22.2, SD = 5.1; in-person M = 22.3, SD = 7.5; $t(25) = 0.1$, $p = 0.96$), Vineland ABC score (telehealth M = 74.8, SD = 13.6; in-person M = 71.7, SD = 10.5; $t(24) = -0.7$, $p =$

0.51), family demographics including socioeconomic status of family income $50,000 and more compared to family incomes less than $50,000 ($X^2(1) = 0.1$), and highest level of caregiver education of caregiver 1 ($X^2(1) = 11.5$) and caregiver 2 ($X^2(1) = 5.5$). The participant's sex assigned at birth was not compared due to the small number of females in the sample.

Research Question 1

Regarding the initial diagnostic impressions provided by clinicians, the assessment teams generally showed agreement. There was agreement on initial diagnostic impressions between in-person and telehealth clinicians in 66.7% of the cases (N = 18). Of these cases with agreement, 37.0% (N = 10) resulted in an ASD diagnostic impression, and 29.6% (N = 8) concluded that an ASD diagnostic impression was not appropriate. In 33.3% of the participants (N= 9), the assessment teams had different initial diagnostic impressions. Among the cases with diagnostic impression disagreement, the majority of participants (77.8%) did not receive an initial diagnostic impression of ASD at their telehealth appointment, whereas they did receive an ASD diagnostic impression at their in-person appointment. Table 6 shows the overall agreement and disagreement of diagnostic impressions across conditions.

Since the in-person ASD diagnostic measure, the ADOS-2, is the gold standard assessment tool for diagnosing individuals with ASD, the final diagnosis for each participant relied on their in-person appointment. However, the clinic encountered complex cases requiring additional information, such as previous reports, additional caregiver questionnaires, and extensive discussions between both assessment teams to

establish a final diagnostic impression for the participant. Two out of the total 27 participants' final diagnoses did not match their in-person assessment diagnostic impression; due to this small number of participants, statistical analyses to examine whether these participants differed meaningfully from others was not feasible.

Table 6

Diagnostic Impressions Across Conditions

Telehealth Diagnostic Impression	In-Person ASD Diagnostic Impression		
	ASD Diagnosis	No ASD Diagnosis	Total
ASD	10	7	17
No ASD	2	8	10
	12	15	27

Independent sample t-tests were used to analyze the impact of participant characteristics on the agreement between telehealth and in-person initial diagnostic impression. No significant differences were observed in participant characteristics between groups where there was an agreement between the initial diagnostic impressions (ASD – ASD and no ASD – no ASD) and those with disagreement (ASD – no ASD and no ASD – ASD). The two groups did not significantly differ based on the participant's chronological age in months at their first appointment (agreement M = 76.2, SD = 38.8; disagreement M = 108.2, SD = 44.5, $t(25) = -1.9$, $p = 0.07$), FSIQ (agreement M = 89.8, SD = 22.3; disagreement M = 82.5, SD = 25.0, $t(20) = 0.7$, $p = 0.49$), Total Behavioral score (agreement M = 2.6, SD = 1.2; disagreement M = 2.8, SD = 0.8, $t(24) = -0.4$, $p = 0.68$), SCQ total scores (agreement M = 23.6, SD = 5.7; disagreement M = 19.6, SD =

7.4, $t(25) = 1.6$, $p = 0.13$), and Vineland ABC score (agreement M = 73.4, SD = 12.2; disagreement M = 72.2, SD = 11.5, $t(24) = 0.2$, $p = 0.81$).

Sensitivity and Specificity

The sensitivity and specificity for TELE-ASD-PEDS and TELE-ASD-KIDS were calculated separately. Among the seven participants who were administered the TELE-ASD-PEDS, or those with limited verbal language, the sensitivity was determined to be 0.83. The specificity value could not be calculated for TELE-ASD-PEDS as no participants fell into the no ASD – no ASD group. Table 7 displays the sensitivity and specificity of the TELE-ASD-PEDS.

Table 7

Sensitivity and Specificity of TELE-ASD-PEDS

	In-Person ASD Diagnostic Impression		
Telehealth Diagnostic Impression	ASD Diagnosis	No ASD Diagnosis	Total
ASD	5	1	6
No ASD	1	0	1
	6	1	7

The sensitivity and specificity were computed for the TELE-ASD-KIDS assessment. Module 2 and Module 3 were combined for this calculation. The sensitivity for TELE-ASD-KIDS was calculated to be 0.40, while the specificity was calculated to be 0.80. Table 8 shows the sensitivity and specificity of TELE-ASD-KIDS.

As expected, the participant's age (in months) at their telehealth appointment significantly differed across the different telehealth ASD protocols. A one-way ANOVA investigated the participant's age between the different telehealth ASD modules. The age

of the participant at their telehealth appointment increased from TELE-ASD-PEDS (N = 7, M = 52.3, SD = 15.0), to TELE-ASD-KIDS Module 2 (N = 6, M = 78.3, SD = 43.3), to TELE-ASD-KIDS Module 3 (N = 14, M = 108.7, SD = 40.5), $F(2, 24) = 5.78$, $p = 0.009$.

Table 8

Sensitivity and Specificity of TELE-ASD-KIDS

Telehealth Diagnostic Impression	In-Person ASD Diagnostic Impression		
	ASD Diagnosis	No ASD Diagnosis	Total
ASD	4	2	6
No ASD	6	8	14
	10	10	20

Clinician Confidence Levels

In terms of clinician confidence levels for initial diagnostic impressions, the prevailing trend was clinicians had some degree of certainty about the initial diagnostic impression at each appointment (refer to Table 9). Specifically, 74.1% of telehealth participants and 92.6% of the in-person participants were rated with some level of certainty (i.e., somewhat certain and completely certain). No initial diagnostic impressions were rated with complete uncertainty. Chi-square test of homogeneity was not completed because the cell count assumption was violated, meaning some cells had less than 5 data points. Therefore, a Fisher's exact test was used to investigate if there were significant differences between the confidence levels of the initial diagnostic impressions between the two assessment styles. The results of the Fisher's exact test ($p = 0.46$) did not indicate a significant relationship between the telehealth and in-person assessment appointments.

Table 9

Clinician Confidence Levels for Diagnostic Impressions

Confidence Level	Telehealth Assessment	In-Person Assessment
Completely Certain	55.6% (N = 15)	74.1% (N = 20)
Somewhat Certain	18.5% (N = 5)	18.5% (N = 5)
Somewhat Uncertain	25.9% (N = 7)	7.4% (N = 2)
Completely Uncertain	0.0% (N = 0)	0.0% (N = 0)

Independent sample t-tests were conducted to determine if there were group differences between participants who were rated with some level of certainty (i.e., somewhat certain, completely certain) to those with some level of uncertainty (i.e., somewhat uncertain, completely uncertain) with respect to the diagnostic impressions from ONLY the telehealth appointment. This analysis was only conducted for telehealth appointments, as it was important to understand which participant characteristics influenced a clinician's confidence for their telehealth diagnostic impression. The mean participant characteristic scores between the certain and uncertain telehealth diagnostic impressions are shown in Table 10.

There was no statistically significant difference between the mean chronological age of participants at their telehealth appointment between the certain (M = 87.4, SD = 42.4) and uncertain (M = 87.1, SD = 47.3) initial diagnostic impressions, $p > 0.05$. There was no statistically significant difference in the average number of previous diagnoses between the certain (M = 1.1, SD = 1.3) and uncertain (M = 1.6, SD = 2.0) diagnostic impression groups, $p > 0.05$. There was no statistically significant difference in the average FSIQ scores between the certain (M = 83.6, SD = 23.5) and uncertain (M = 99.2, SD = 18.2) diagnostic impression groups, $p > 0.05$. There was no statistically significant

difference in the mean ASD telehealth total Likert score between the certain (M = 16.3, SD = 2.4) and uncertain (M = 16.7, SD = 2.3) diagnostic impression groups, $p > 0.05$. There was no statistically significant difference in the mean Vineland ABC between the certain (M = 70.8, SD = 11.7) and uncertain (M = 78.9, SD = 10.3) diagnostic impression groups, $p > 0.05$. There was no statistically significant difference in the mean Total Other Behavior score between the certain (M = 1.1, SD = 1.3) and uncertain (M = 1.6, SD = 2.0) diagnostic impression groups at the telehealth appointment, $p > 0.05$.

There was a statistically significant difference in the mean SCQ total score between the certain (M = 24.3, SD = 6.0) and uncertain (M = 16.6, SD = 4.3) diagnostic impression groups at the telehealth appointment, $t(25) = -3.1$, $p = 0.005$. This indicates that participants with higher SCQ scores, or those with more reported ASD related behaviors by caregivers, had more clinician certainty with their telehealth diagnostic impression, as compared to those with lower SCQ total scores.

Chi-square test of homogeneity was not completed because the cell count assumption was violated, meaning some cells had less than 5 data points. Therefore, a Fisher's exact test investigated the clinician's confidence of telehealth diagnostic impressions and participant's characteristics. The results of the Fisher's exact test (p = 0.41) did not indicate a significant relationship between those provided a final ASD diagnosis between certain (66.7%) and uncertain (33.3%) diagnostic impressions.

Table 10

Mean Participant Characteristics Based on Telehealth Clinician Confidence Levels

Child Characteristics	Completely and Somewhat Certain N = 20	Completely and Somewhat Uncertain N = 7
Sex of child (% males)	85.0%	71.4%
Average chronological age at telehealth appointment (months)	87.4 (42.4)	87.1 (47.3)
Average number of previous diagnoses	1.1 (1.3)	1.6 (2.0)
Average Full-Scale IQ[a]	83.6 (23.5)	99.2 (18.2)
Telehealth Total Likert Score	16.2 (2.4)	16.7 (2.3)
Vineland-3 Adaptive Behavior Composite	70.8 (11.7)	78.9 (10.3)
SCQ Total Score	**24.3 (6.0)**	**16.6 (4.3)**
Total Other Behavior Score	2.6 (1.0)	2.8 (1.3)
ASD Final Diagnosis	66.7%	33.3%

[a]*Full Scale IQ score was calculated either by the WPPSI-IV or WASI-II, which were administered based on the child's age.*
Note. Bold values indicate $p < .01$

Research Question 2

A paired sample t-test analyzed the difference in caregivers' social validity ratings between telehealth and in-person appointments. Out of the seven statements, two statements had statistically significant differences. When families were asked if they felt the assessment team at each respective appointment saw an accurate representation of their child, there was a statistically significant difference between telehealth (M = 3.4, SD = 1.2) and in-person (M = 4.2, SD = 1.0), $t(13) = 2.8$, $p = 0.02$. This result indicated that caregivers perceived it was more difficult for clinicians to obtain important information about their child when assessed through telehealth then in-person. Additionally, when families were asked if they would participate in that assessment style again, there was a statistically significant difference between telehealth (M = 3.9, SD = 0.7) and in-person

(M = 4.6, SD = 0.5), $t(13) = 3.2$, $p = 0.006$. This exemplified that caregivers reported they were more likely to participant in in-person ASD evaluation than telehealth ASD appointment again.

Table 11 shows caregivers' mean social validity scores for each social validity statement. Only caregivers who responded to both the telehealth and in-person social validity questionnaires were included in this analysis (N = 14). The bolded items indicate statistically significant differences between the two groups.

Not all social validity statements were relevant for each assessment process. Therefore, certain statements were not assessed using the paired sample t-test analysis. For example, out of the 23 families who completed the in-person social validity scale, five caregivers reported that they agreed or strongly agreed that it was inconvenient to drive to the clinic (21.7%). The average commute time for families was 43.4 minutes, ranging from 3.0 to 100.0 minutes. Of the 14 families who completed the social validity questionnaire, most families previously participated in telehealth appointments, N = 11 (78.6%). Additionally, most caregivers disagreed with the statement, “I needed to learn a lot of things about technology before I could start this telehealth assessment” (N = 13, 92.9%). Most of the families agreed with the following two statements: “The technology used for this assessment worked well for me” (N = 13, 92.9%) and “I felt that my child’s telehealth assessment was just as private as in-person visit would be” (N =13, 92.9%).

Table 11

Mean Scores for Social Validity Questionnaire Statements

Statements	Telehealth N = 14	In-Person N = 14
1. I thought the appointment was easy.	4.5 (0.5)	4.6 (0.6)
2. Most people would find this assessment process easy to follow.	4.4 (0.6)	4.6 (0.5)
3. I was able to communicate my concerns to the assessment team.	4.7 (0.5)	4.8 (0.4)
4. I felt that the assessor was able to collect important information about my child.	4.1 (0.8)	4.6 (0.6)
5. The assessment team saw an accurate representation of my child's behavior.	**3.4 (1.2)**	**4.2 (1.0)**
6. I would participate in this assessment style again.	**3.9 (0.7)**	**4.6 (0.5)**
7. Overall, I am satisfied with my assessment experience.	4.4 (0.7)	4.6 (0.5)

Note: Only families who completed both telehealth and in-person social validity questionnaires are represented in this table. Bold values indicate $p < 0.05$.

Caregivers were also asked to compare their experiences from the telehealth and in-person appointments. The questions asked caregivers if there was a preferred assessment style. Caregivers responded to each statement with one of the following options: Telehealth, In-Person, Both, or Neither. Most caregivers reported that in-person appointments were easier for their child (57.1%), and caregivers would have chosen an in-person appointment for their child's ASD evaluation (85.7%). Percentages were calculated based on each statement's responses, which are provided in Table 12. Due to the small number of responses, only descriptive statistics are provided.

Table 12

Percentages of Caregiver Preferences Based on Assessment Style

Statement	Assessment Style (N = 13)			
	Telehealth (%)	**In-Person (%)**	**Both (%)**	**Neither (%)**
1. Which assessment felt more private to you?	7.1%	28.6%	64.3%	0.0%
2. Which assessment was more convenient for you?	35.7%	21.4%	42.9%	0.0%
3. Which assessment were you more comfortable with?	14.3%	21.4%	64.3%	0.0%
4. Which assessment helped you most effectively communicate your concerns to the clinical team?	7.1%	42.9%	50.0%	0.0%
5. Which assessment process was easier for you to participate in?	7.1%	28.6%	64.3%	0.0%
6. Which assessment process was easier for your child to participate in?	7.1%	57.1%	35.7%	0.0%
7. If you could only choose one (telehealth or in person), which process would you have participated in for your child's autism evaluation?	0.0%	85.7%	14.3%	0.0%

Child participants with fluent speech were asked about their assessment experience. Of the 11 participants provided the questions, 80.0% liked participating over Zoom or at home, and 81.8% liked coming to the clinic. Child participant responses are displayed in Table 13.

Table 13

Child Participants' Perception of Assessment Styles

	Child Responses (N = 11)		
Statement	**Strongly Disagree/ Disagree (%)**	**Neither Agree or Disagree (%)**	**Agree/ Strongly Agree (%)**
1. I liked coming to the clinic	18.2%	0.0%	81.8%
2. I liked participating over Zoom/at home.	0.0%	20.0%	80.0%
3. It did not matter to me where you saw me.	10.0%	50.0%	40.0%

Research Question 2A

The second part of this research question explored participant and family characteristics associated with caregivers' telehealth social validity ratings. The social validity telehealth questions were averaged to produce a Total Social Validity Score. The Total Social Validity Score was created by averaging all the social validity responses. One social validity statement was reversed coded before averaging the social validity statements as it was the only negatively phrased statement. The Total Social Validity score was correlated to participant and family characteristics.

Spearman's rank-order correlations were conducted to analyze the relationships between the telehealth Total Social Validity score and the following variables: adaptive skills (Vineland ABC), full-scale IQ (FSIQ), Total Other Behavior score, SCQ total score, Telehealth Total Likert score, Telehealth Clinician Confidence level, and travel time to clinic (N = 14). Due to violations to the assumption of linear relationships for a Pearson's correlation, a non-parametric correlation was utilized. It should be noted that

two FSIQ scores and one Vineland ABC score were missing from this analysis. No statistically significant correlations were found between the Total Social Validity scores and the following variables: Adaptive skills (Vineland ABC), full scale IQ, Total Other Behavior score, caregiver reported ASD behaviors (SCQ), Telehealth ASD assessment Likert total score, telehealth clinician confidence score on initial diagnostic impressions, and the clinic travel time. Correlation coefficients can be found in Table 14.

A point-biserial correlation analyzed the relationship between the Total Social Validity score and dichotomous family and child characteristics, including family income and the participant's race and ethnic background. The child's sex was not included in this analysis due to the limited number of females assigned at birth included in the data. The family's income was divided into two groups: $50,000 or less and greater than $50,000. The child's racial and ethnic background was separated into two groups: non-Hispanic/Latinx White origin and of Hispanic/Latinx origin. The assumption of normality was violated for family income $50,000 and less as assessed by Shapiro-Wilk's test, $p = 0.03$. The assumption of equal variances was met for this variable, as assessed by Levene's Test for Equality of Variances, $p > 0.05$. When data is mildly shewed or mildly not normally distributed and the variances are equal, the point-biserial correlation results can still be interpreted (NCSS Statistical Software, n.d.). One family income was not reported by a family and therefore it was missing for this analysis. No statistically significant correlations were found between the Total Social Validity scores and other family and child characteristics. Correlation coefficients are displayed in Table 14.

Table 14

Correlation Coefficients with Telehealth Social Validity Total Score

Variable	Correlation Coefficient with Telehealth Total Social Validity Score
Vineland ABC	0.36
FSIQ	-0.51
Total Other Behavior Score	0.12
SCQ Total Score	-0.10
Telehealth Total Likert Score	0.23
Telehealth Clinician Confidence Score	-0.14
Travel Time to Clinic	0.00
Family Income	-0.36
Child Race/Ethnicity	0.30

Note. Vineland ABC = Vineland Adaptive Behavior Composite score; FSIQ = Full Scale IQ score, SCQ = Social Communication Questionnaire.

Discussion

The present study investigated the agreement between clinician diagnostic impressions obtained from telehealth ASD evaluations and the gold standard in-person ASD evaluation, the ADOS-2. While telehealth ASD assessment tools are shorter and require fewer behavioral codes, it was important to determine whether telehealth ASD evaluations accurately captured ASD behaviors. Furthermore, the study aimed to identify child characteristics that were conducive to telehealth ASD evaluations, as well as participant characteristics that impacted a clinicians' confidence for their telehealth diagnostic impression. This study also examined the family's experience and investigated if family characteristics influenced their telehealth experience. Overall, this study provided promising results to the continued use of telehealth ASD assessments as there was high agreement between telehealth and in-person diagnostic impressions. Caregivers

also reported positive telehealth experiences without the need to learn new skills to participate virtually.

Research Question 1

This study exemplified how modifying in-person ASD evaluations to virtual methods (e.g., providing cognitive and ASD evaluations virtually with caregiver questionnaires and interview) resulted in high agreement (66.7%) between in-person and telehealth diagnostic impressions. This finding aligned with prior research, which found high diagnostic impression agreement between in-person and telehealth evaluations. When comparing in-person to telehealth ASD assessment diagnostic results, Corona and colleagues (2023) had 92% agreement, Reese et al. (2013) had 86% agreement, and Smith and colleagues (2017) had 88.2% agreement. It should be noted that the studies previously cited consisted of younger participants compared to the current study. The other research teams' younger participants may have had an impact of the higher levels of agreement compared to this paper, which included older and more verbal participants.

Past studies produced mixed findings concerning participant characteristics that influenced agreement and disagreement between telehealth and in-person diagnostic impressions. Although this study did not yield statistically significant differences between participant characteristics and the degree of agreement/disagreement between telehealth and in-person, it is anticipated that with a larger sample, some differences might emerge. Due to the small sample size of this study, dividing the sample into three groups was not possible. Future analyses should investigate participant characteristics between ASD agreement (ASD – ASD), no ASD agreement (no ASD – no ASD) and

diagnostic impression disagreement (no ASD – ASD and ASD – no ASD) using ANOVA analysis. Expected differences between the groups would include the participant's age, as the preliminary results indicated higher sensitivity for those administered the TELE-ASD-PEDS (M = 52.3 months old) compared to TELE-ASD-KIDS (Module 2 M = 78.3 months old, Module 3 M = 108.7 months old). Other participant characteristics hypothesized to differ between diagnostic agreement/disagreement include Total Other Behavior score, adaptive skills, and cognitive functioning, as suggested by the literature (Gibb et al., 2021; Kryszak et al., 2022; Wagner et al., 2021a; Wagner et al., 2022). It would be expected to see higher diagnostic impression agreement for participants who were younger, lower in cognitive functioning, lower in adaptive skills, and those having fewer other non-related ASD behaviors (e.g., anxiety, tantrums, hyperactivity) as implied by the literature (Kryszak et al., 2022; Wagner et al., 2021a; Wagner et al., 2022).

In the instances of diagnostic impression disagreement, most cases in this study resulted with telehealth appointments initially indicating the participant did not meet criteria for ASD, whereas the in-person appointment concluded the participant did meet ASD criteria. This highlights possible factors that make it difficult to observe ASD behaviors virtually, either due to the subtleness of behaviors or due to technology barriers (e.g., poor camera angles, poor audio quality; Corona et al. 2021; Gibbs et al., 2021; Ludwig et al., 2021; Talbott et al., 2022; Wagner et al., 2021a). Other researchers have noted challenges in observing subtle behaviors, such as eye contact, gestures, RRBs, through video conferencing (Corona et al., 2021; Gibbs et al., 2021; Jones et al., 2022). For example, a Module 3 teenage participant assessed as part of this study, was rated as

having great eye contact during their telehealth assessment due to consistent eye contact with their computer screen. However, at the in-person appointment, the participant rarely provided eye contact with the assessment team. This one example exemplifies how subtle behaviors may be misinterpreted, exaggerated, or missed during telehealth assessments.

Sensitivity and Specificity

As a further measure of diagnostic impressions, sensitivity and specificity were calculated. Although the analysis involved a small number of participants, previous studies have computed sensitivity and specificity of telehealth assessments with participant numbers as low as five (Nazneen et al., 2015) or 17 cases (Reese et al., 2015). For comparison, the gold standard ASD assessment tool, ADOS-2, has a sensitivity ranging from 0.82 to 0.90 for Modules 1, 2, and 3. The specificity of the ADOS-2 for the same modules ranges from 0.62 to 0.90 (Dorlack et al., 2018). Previous studies also calculated the sensitivity and specificity for TELE-ASD-PEDS to be 0.80 and 1.00 respectively (Jones et al., 2022). In the current study, TELE-ASD-PEDS sensitivity was calculated at 0.83, indicating a high degree of concluding a true ASD diagnosis. Currently, the specificity was not calculated due to insufficient participants in the no ASD agreement group (i.e., no ASD – no ASD). As for TELE-ASD-KIDS, the sensitivity and specificity were calculated as 0.36 and 0.80, respectively. These preliminary results indicated higher accuracy in correctly identifying ASD through telehealth ASD assessments with younger and less verbal individuals. Thus, further supporting that telehealth assessments may be better at identifying ASD in younger individuals (Kryszak et al., 2022; Wagner et al., 2021a; Wagner et al., 2022).

Participant Characteristics Influence on Telehealth Diagnostic Confidence Levels

The literature has emphasized mixed results for factors that influenced clinicians' confidence in telehealth diagnostic impressions. Although previous studies corroborated high clinician confidence in diagnostic impressions from telehealth ASD evaluations, there were variables that impacted a clinician's confidence level (Corona et al., 2022). Clinicians identified participant characteristics, such as the participant's age, cognitive functioning, language abilities, and adaptive skills, as influential (Gibbs et al., 2021; Kryszak et al., 2022; Wagner et al., 2021a; Wagner et al., 2022). This emphasized how some key diagnostic features, such as eye contact, RRBs were potentially out of the camera's view, or other more subtle behaviors may not have been fully captured through telehealth processes. Missing these key behaviors could influence the diagnostic impressions and behavioral codes concluded from telehealth assessments (Kryszak et al., 2022).

It is important to note that this study found more clinicians rated their confidence levels as completely certain when assessing a participant in-person compared to telehealth. This may highlight the need to change some of the features related to telehealth ASD evaluations to help increase the clinicians' confidence in their diagnostic impressions. Such alterations could include better camera angles (e.g., ensuring the participant is entirely in view) and potentially reducing environmental distractions that could interfere with the assessment (e.g., removing other people not involved, eliminating distracting objects). To assist with these changes, clinicians and caregivers may need to schedule a meeting before the appointment to help with these common areas of difficulty

or barriers to telehealth appointments. Such suggestions were raised from caregivers in Talbott and colleagues study (2022).

When participant characteristics were investigated compared to telehealth diagnostic impression confidence levels in this study, the only significant difference found between those who rated their diagnostic impression with certainty and uncertainty were for SCQ total scores. The higher scores on the SCQ were associated with more caregiver-reported ASD related behaviors. This result showed that more caregiver-reported ASD related behaviors was related to higher certainty with the clinician's diagnostic impression from their telehealth assessment. This might indicate that those participants with higher SCQ totals display more obvious ASD-like behaviors and, therefore, made it easier for trained professionals to provide a more confident diagnostic impression for that individual. This further supports how individuals with more subtle behaviors might be harder to diagnosis when assessed through telehealth.

With a larger sample recruited for this study, significant differences in other participant characteristics, such as age, adaptive abilities, cognitive functioning, and other non-related ASD behaviors that impacted the evaluation might also emerge. With a larger sample pool, these participant characteristics may influence the success and accuracy of telehealth diagnostic impressions and confidence levels. Additionally, with a larger sample size it would be possible to determine whether participant characteristics, such as language ability or age, play a significant role in clinician's confidence in making telehealth diagnostic impressions.

Social Validity Responses

Caregivers and participants were asked about their experience regarding the easiness of the appointment, the ability to communicate with clinicians, the clinicians' ability to collect information, and their satisfaction with their assessment experience; which all did not significantly differ between the two assessment styles. These results aligned with previous studies. Across the literature, caregivers were just as satisfied with telehealth ASD evaluations as in-person appointments, were able to effectively communicate their concerns with clinicians, and develop rapport virtually (Corona et al., 2021; Gibbs et al., 2021; Pompa-Craven et al., 2022; Reese et al., 2013; Reisinger et al., 2022; Talbott et al., 2022).

Overall, caregivers reported higher social validity scores on items favoring in-person appointments, as corroborated by past studies (Jones et al., 2022; Talbott et al., 2022). Although there were benefits for telehealth appointments, such as reduced travel time to appointments, and rapport was not impacted, caregivers might have preferred in-person appointments because they believed the clinician missed key information or did not observe notable behaviors on the telehealth screen. As noted by other studies, the need to learn more technology skills for both the caregiver and child could have impacted the families' assessment preference for in-person evaluations, such as learning about the correct camera angles and listening to live coaching (Talbott et al., 2022).

The preliminary findings of this study, as well as other literature on this topic, highlight the need to help educate caregivers and participants on the technology and filming protocols. A possible solution is to schedule an appointment with caregivers

before the evaluation appointment to review materials needed for the appointment, provide information about the structure of the assessment, and practice filming (e.g., camera angle, sitting area, reducing distractions; Talbott et. al., 2022). It is important to address these potential influencing factors to make the family's telehealth ASD experience just as comfortable as an in-person appointment.

Youth participants were also asked about their experience. The participants did not have a strong preference for either assessment style. This indicates individuals being assessed are comfortable with evaluations provided in-person and virtually. This report also suggests that the participants did not have a preferred assessment style.

Researchers in previous studies have identified various factors that influenced a family's telehealth experience, such as the family's socioeconomic status, the child's assigned sex at birth, and child's adaptive functioning (Gibbs et al., 2021; Kryszak et al., 2022; Pomales-Ramos et al., 2023; Reisinger et al., 2022). In this preliminary analysis, no statistically significant correlations were found among family and participant variables and the caregiver's telehealth experience as reported through social validity questionnaires. This is potentially due to a small sample size. With more social validity responses, it can be suspected that the child's adaptive, cognitive functioning, and other behaviors not related to ASD, as well as the number of people in the household, may become statistically significant as it is reported to be impactful to telehealth experiences. As such, researchers have found common barriers to telehealth services include challenging behaviors (e.g., tantrums, elopement), limited language abilities, and lower adaptive skills (Pomales-Ramos et al., 2023; Reisinger et al., 2022). Caregivers have

reported factors that impacted their telehealth experiences were distractions, technology difficulties, and challenges engaging their child in the assessment activities (Talbott et al., 2022).

Limitations

This preliminary study included a subset of the total participants of this research project. As such, more participants can further determine if there are differences between diagnostic impressions, clinicians' confidence levels for telehealth diagnostic impressions, as well as in the specificity and the sensitivity for the telehealth ASD assessments. A larger sample will hopefully include more social validity responses to help researchers and clinicians understand the barriers and challenges of telehealth ASD evaluations. From the current participants included in this study, some preliminary barriers included the easiness of the telehealth ASD assessment for the child (e.g., operating Zoom, camera angles) and accurately capturing the participant's concerning behaviors. This input can help change and modify the telehealth experience for families to make them feel as comfortable as if the evaluation was in-person.

A current limitation of this study is the sample used. The current sample included families currently seeking an evaluation from the SEARCH center **and** who gave consent to participate in the study. Notably, this study did not include families who declined to participate. It is possible that the families who did consent may differ in certain ways from those who chose not to. This issue often piques the interest of researchers, as they seek to understand why some individuals opt to participate while other do not. It should be acknowledged that we did not capture data from all families seeking an ASD

evaluation in the community. It is important to note that the families who declined to participate received a full ASD evaluation in-person with a full report following the completion of their appointment.

Strengths of this Study

Despite this being a preliminary study, the results added to a growing literature on ASD telehealth assessments. The study showed promising results that further supported the use of telehealth ASD assessments. A strength of this study was that the participants were from the local community where many residents were of lower socioeconomic status and lived in a "resource desert." The families in this community are ideally the families who would benefit most from the continuation of telehealth ASD evaluations (e.g., less travel to clinics, equitable access to knowledgeable assessors). The caregivers reported overall positive experiences from their telehealth appointments.

Additionally, the telehealth assessments indicated reliable identification of ASD in participants with limited to phrase speech. These participants were typically younger and therefore one might argue that widespread use of telehealth could help identify children at a younger age. As previously discussed, diagnosing a child at a younger age leads to earlier provision of intervention to young children. Earlier intervention has shown to provide a more favorable prognosis for those individuals.

Another strength of this study was its diversity of participants. First, this study recruited a wide age range of participants. The average age of participants was roughly seven years old, with the majority of participants administered a more verbal assessment. Additionally the majority of our participants identified as being Hispanic/Latinx. Unlike

many past studies, including studies asking caregivers about their telehealth experiences, which predominately included younger participants with less verbal abilities and White participants (Corona et al., 2022; Pomales-Ramos et al., 2023; Reisinger et al., 2022; Smith et al., 2017; Stavropoulos et al., 2022; Talbott et al., 2022; Wagner et al., 2021a). Additionally, the participants were local to the center, situated in a medically underserved area. The clinic also serves families living in rural areas, who often travel long distances to access medical services, such as ASD evaluations.

Implications

In-person evaluations are not always feasible and accessible for all families. Therefore, developing alternative evaluations for all families, even after the COVID-19 pandemic, is necessary to be sure all children are diagnosed in a timely manner and receive appropriate early interventions. The high level of diagnostic agreement between in-person and telehealth ASD evaluations is promising and promotes the continued use of telehealth ASD assessments, especially for families who find this method more convenient.

Future studies need to further investigate child and family characteristics that may influence diagnostic impressions and diagnostic confidence when using telehealth ASD assessments, along with continual assessment of the diagnostic accuracy of telehealth ASD evaluations. Caregiver input should be collected as it will continue to shape the telehealth ASD evaluation, to help families feel more comfortable with this quicker and shorter assessment style. Furthermore, telehealth can help change how ASD evaluations

are administered in the future and help provide choices for families seeking an ASD assessment.

References

Achenbach, T. M., & Rescorla, L. A. (2001). *Manual for the ASEBA school-age forms & profiles: An integrated system of multi-informant assessment*. University of Vermont.

Alfuraydan, M., Croxall, J., Hurt, L., Kerr, M., & Brophy, S. (2020). Use of telehealth for facilitating the diagnostic assessment of autism spectrum disorder (ASD): A scoping review. *PLoS ONE, 15*(7), Article e0236415. https://doi.org/10.1371/journal.pone.0236415

Altman, D. G., & Bland, J. M. (1994). Diagnostic tests 1: Sensitivity and specificity. *British Medical Journal, 308*(6943), 1552. https://doi.org/10.1136/bmj.308.6943.1552

Ameis, S. H., Lai, M. C., Mulsant, B. H., & Szatmari, P. (2020). Coping, fostering resilience, and driving care innovation for autistic people and their families during the COVID-19 pandemic and beyond. *Molecular Autism, 11*(1), 1-9. https://doi.org/10.1186/s13229-020-00365-y

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

Antezana, L., Scarpa, A., Valdespino, A., Albright, J., & Richey, J. A. (2017). Rural trends in diagnosis and services for autism spectrum disorder. *Frontiers in Psychology*, *8*, 1-5. https://doi.org/10.3389/fpsyg.2017.00590

Avlund, S. H., Thomsen, P. H., Schendel, D., Jørgensen, M., Carlsen, A. H., & Clausen, L. (2021). Factors associated with a delayed autism spectrum disorder diagnosis in children previously assessed on suspicion of autism. *Journal of Autism and Developmental Disorders*, *51*(11), 3843-3856. https://doi.org/10.1007/s10803-020-04849-x

Baratloo, A., Hosseini, M., Negida, A., & El Ashal, G. (2015). Part 1: Simple definition and calculation of accuracy, sensitivity and specificity. *Emergency, 3*(2), 48-49.

Baweja, R., Brown, S. L., Edwards, E. M., & Murray, M. J. (2022). COVID-19 pandemic and impact on patients with autism spectrum disorder. *Journal of Autism and Developmental Disorders, 52*(1), 473-482. https://doi.org/10.1007/s10803-021-04950-9

Beckett, M. K., & Morrison, P. A. (2010). Assessing the need for a new medical school:

A case study in applied demography. *Population Research and Policy Review*, *29*, 19-32. https://doi.org/10.1007/s11113-009-9145-6

Berger, N. I., Wainer, A. L., Kuhn, J., Bearss, K., Attar, S., Carter, A. S., Ibanez, L.V., Ingersoll, B. R., Neiderman, H., Scott, S., & Stone, W. L. (2022). Characterizing available tools for synchronous virtual assessment of toddlers with suspected autism spectrum disorder: A brief report. *Journal of Autism and Developmental Disorders*, *52*(1), 423-434. https://doi.org/10.1007/s10803-021-04911-2

Best, A. L., Fletcher, F. E., Kadono, M., & Warren, R. C. (2021). Institutional distrust among African Americans and building trustworthiness in the COVID-19 response: Implications for ethical public health practice. *Journal of Health Care for the Poor and Underserved, 32*(1), 90-98. https://doi.org/10.1353/hpu.2021.0010

Bishop-Fitzpatrick, L., & Kind, A. J. H. (2017). A scoping review of health disparities in autism spectrum disorder. *Journal of Autism and Developmental Disorders, 47*(11), 3380–3391. https://doi.org/10.1007/s10803-017-3251-9

Bivarchi, F. A., Kehyayan, V., & Al-Kohji, S. M. (2021). Barriers to the early detection and intervention of children with autism spectrum disorders: A literature. *Journal of Nursing Education and Practice*, *11*(11), 72-80. https://doi.org/10.5430/jnep.v11n11p72

Blacher, J., Cohen, S. R., & Azad, G. (2014). In the eye of the beholder: Reports of autism symptoms by Anglo and Latino mothers. *Research in Autism Spectrum Disorders*, *8*(12), 1648-1656. https://doi.org/10.1016/j.rasd.2014.08.017

Blacher, J., Stavropoulos, K., & Bolourian, Y. (2019). Anglo-Latino differences in parental concerns and service inequities for children at risk of ASD. *Autism*, *23*(6), 1554-1562. DOI:10.1177/1362361318818327

Campos-Castillo, C., & Anthony, D. (2021). Racial and ethnic differences in self-reported telehealth use during the COVID-19 pandemic: A secondary analysis of a US survey of internet users from late March. *Journal of the American Medical Informatics Association, 28*(1), 119-125. https://doi.org/10.1093/jamia/ocaa221

Center for Disease Control and Prevention [CDC]. (2020, March 2). *Data & statistics on autism spectrum disorder.* https://www.cdc.gov/ncbddd/autism/data.html

Chu, K. (1999). An introduction to sensitivity, specificity, predictive values and likelihood ratios. *Emergency Medicine, 11*(3), 175-181. https://doi.org/10.1046/j.1442-2026.1999.00041.x

Cidav, Z., Munson, J., Estes, A., Dawson, G., Rogers, S., & Mandell, D. (2017). Cost offset associated with early start Denver model for children with autism. *Journal of the American Academy of Child & Adolescent Psychiatry, 56*(9), 777-783. https://doi.org/10.1016/j.jaac.2017.06.007

Constantino, J. N. (2012). *Social Responsiveness Scale, Second Edition (SRS-2)*. Western Psychological Services.

Corona, L., Hine, J., Nicholson, A., Stone, C., Swanson, A., Wade, J., Wagner, L., Weitlauf, A., & Warren, Z. (2020). TELE-ASD-PEDS: A Telemedicine-based ASD Evaluation Tool for Toddlers and Young Children. Vanderbilt University Medical Center. https://vkc.vumc.org/vkc/triad/tele-asd-peds

Corona, L. L., Wagner, L., Hooper, M., Weitlauf, A., Foster, T. E., Hine, J., Miceli, A., Nicholson, A., Stone, C., Vehorn, A., & Warren, Z. (2023). A randomized trial of the accuracy of novel telehealth instruments for the assessment of autism in toddlers. *Journal of Autism and Developmental Disorders*, 1-12. https://doi.org/10.1007/s10803-023-05908-9

Corona, L. L., Weitlauf, A. S., Hine, J., Berman, A., Miceli, A., Nicholson, A., Stone, C., Broderick, N., Francis, S., Pablo Juárez, A., Vehorn, A., Wagner, L., & Warren, Z. (2021). Parent perceptions of caregiver-mediated telemedicine tools for assessing autism risk in toddlers. *Journal of Autism and Developmental Disorders*, *51*(2), 476-486. https://doi.org/10.1007/s10803-020-04554-9

Daniels, A. M., & Mandell, D. S. (2014). Explaining differences in age at autism spectrum disorder diagnosis: A critical review. *Autism*, *18*(5), 583-597. https://doi.org/10.1177/1362361313480277

de Winter, J. F. C., & Dodou, D. (2010). Five-point likert items: T test versus mann-whitney-wilcoxon. *Practical Assessment, Research, and Evaluation, 15*, 1-16. https://doi.org/10.7275/bj1p-ts64

Dorlack, T. P., Myers, O. B., & Kodituwakku, P. W. (2018). A comparative analysis of the ADOS-G and ADOS-2 algorithms: Preliminary findings. *Journal of Autism and Developmental Disorders*, *48*(6), 2078-2089. https://doi.org/10.1007/s10803-018-3475-3

Dow, D., Holbrook, A., Toolan, C., McDonald, N., Sterrett, K., Rosen, N., Hyun Kim, S. & Lord, C. (2021). The brief observation of symptoms of autism (BOSA): Development of a new adapted assessment measure for remote telehealth administration through COVID-19 and beyond. *Journal of Autism and Developmental Disorders*, *52*, 1-12. https://doi.org/10.1007/s10803-021-05395-w

Durkin, M. S., Maenner, M. J., Meaney, F. J., Levy, S. E., DiGuiseppi, C., Nicholas, J. S., Kirby, R. S., Pinto-Martin, J. A., & Schieve, L. A. (2010). Socioeconomic inequality in the prevalence of autism spectrum disorder: Evidence from a U.S. cross-sectional study. *PloS one, 5*(7), Article e11551. https://doi.org/10.1371/journal.pone.0011551

Elder, J. H., Brasher, S., & Alexander, B. (2016). Identifying the barriers to early diagnosis and treatment in underserved individuals with autism spectrum disorders (ASD) and their families: A qualitative study. *Issues in Mental Health Nursing*, *37*(6), 412-420. https://doi.org/10.3109/01612840.2016.1153174

Elliott, C. D. (2007). *Differential Ability Scales (2nd ed.)*. Pearson.

Eshraghi, A. A., Li, C., Alessandri, M., Messinger, D. S., Eshraghi, R. S., Mittal, R., & Armstrong, F. D. (2020). COVID-19: Overcoming the challenges faced by individuals with autism and their families. *The Lancet Psychiatry, 7*(5), 481-483. https://doi.org/10.1016/S2215-0366(20)30197-8

Fountain, C., King, M. D., & Bearman, P. S. (2011). Age of diagnosis for autism: Individual and community factors across 10 birth cohorts. *Journal of Epidemiology & Community Health*, *65*(6), 503-510. https://doi.org/10.1136/jech.2009.104588

Frank, H. E., Grumbach, N. M., Conrad, S. M., Wheeler, J., & Wolff, J. (2021). Mental health services in primary care: Evidence for the feasibility of telehealth during the COVID-19 pandemic. *Journal of Affective Disorders Reports*, *5*, Article 100146. https://doi.org/10.1016/j.jadr.2021.100146

Franke, T. M., Ho, T., & Christie, C. A. (2012). The chi-square test: Often used and more often misinterpreted. *American Journal of Evaluation*, *33*(3), 448-458. https://doi.org/10.1177/1098214011426594

Gibbs, V., Cai, R. Y., Aldridge, F., & Wong, M. (2021). Autism assessment via telehealth during the Covid 19 pandemic: Experiences and perspectives of autistic adults, parents/carers and clinicians. *Research in Autism Spectrum Disorders*, *88*, Article 101859. https://doi.org/10.1016/j.rasd.2021.101859

Gordon-Lipkin, E., Foster, J., & Peacock, G. (2016). Whittling down the wait time: Exploring models to minimize the delay from initial concern to diagnosis and treatment of autism spectrum disorder. *Pediatric Clinics*, *63*(5), 851-859. https://doi.org/10.1016/j.pcl.2016.06.007

Gotham, K., Pickles, A., & Lord, C. (2012). Trajectories of autism severity in children

using standardized ADOS scores. *Pediatrics*, *130*(5), e1278-e1284. https://doi.org/10.1542/peds.2011-3668

Gresham, F. M., & Elliot, S. N. (2008). *Social Skills Improvement System Rating Scales.* NCS Pearson.

Grzadzinski, R., Carr, T., Colombi, C., McGuire, K., Dufek, S., Pickles, A., Lord, C. (2016). Measuring changes in social communication behaviors: Preliminary development of the brief observation of social communication change (BOSCC). *Journal of Autism and Developmental Disorders, 46*(7), 2464–2479. https://doi.org/10.1007/s1080 3-016-2782-9

Gupta, S., Simon, K. I., & Wing, C. (2020). Mandated and voluntary social distancing during the covid-19 epidemic: A review. *Brookings Papers on Economic Activity, 2020*(2), 269-326. https://doi.org/10.3386/w28139

Guthrie, W., Swineford, L. B., Nottke, C., & Wetherby, A. M. (2013). Early diagnosis of autism spectrum disorder: Stability and change in clinical diagnosis and symptom presentation. *Journal of Child Psychology and Psychiatry*, *54*(5), 582-590. https://doi.org/10.1111/jcpp.12008

Havlicek, L. L., & Peterson, N. L. (1976). Robustness of the Pearson correlation against violations of assumptions. *Perceptual and Motor Skills*, *43*(3), 1319-1334. https://doi.org/10.2466%2Fpms.1976.43.3f.1319

Horlin, C., Falkmer, M., Parsons, R., Albrecht, M. A., & Falkmer, T. (2014). The cost of autism spectrum disorders. *PloS One, 9*(9), Article e106552. https://doi.org/10.1371/journal.pone.0106552

Hyman, S. L., Levy, S. E., Myers, S. M., Council on Children with Disabilities, Section on Developmental and Behavioral Pediatrics; Kuo, D. Z., Apkon, S., Davidson, L. F., Ellerbeck, K. A., Foster, J. E. A., Noritz, G. H., O'Connor Leppert, M., Saunders, B. S., Stille, C., Yin, L., Weitzamn, C. C., Omer Childers, D., Levine, J. M., Peralta-Carcelen, A. M., Poon, J. K., … Bridgemohan, C. (2020). Identification, evaluation, and management of children with autism spectrum disorder. *Pediatrics*, *145*(1), 1-5. https://doi.org/10.1542/peds.2019-3447

Institute., C. H. C. (2020). 2020 SocioNeeds Index. Retrieved from http://www.shaperivco.org/index.php?module=indicators&controller=index&action=socioneeds.

Isensee, C., Schmid, B., Marschik, P. B., Zhang, D., & Poustka, L. (2022). Impact of

COVID-19 pandemic on families living with autism: An online survey. *Research in Developmental Disabilities, 129*, Article 104307. https://doi.org/10.1016/j.ridd.2022.104307

Jang, J., White, S. P., Esler, A. N., Kim, S. H., Klaiman, C., Megerian, J. T., Morse, A., Nadler, C., & Kanne, S. M. (2022). Diagnostic evaluations of autism spectrum disorder during the COVID-19 pandemic. *Journal of Autism and Developmental Disorders*, *52*(2), 962-973. https://doi.org/10.1007/s10803-021-04960-7

Jones, M. K., Zellner, M. A., Hobson, A. N., Levin, A., & Roberts, M. Y. (2022). Understanding caregiver satisfaction with a telediagnostic assessment of autism spectrum disorder. *American Journal of Speech-Language Pathology, 31*(2), 982-990. https://doi.org/10.1044/2021_AJSLP-21-00139

Katzell, R. A., & Cureton, E. E. (1947). Biserial correlation and prediction. *The Journal of Psychology*, *24*(2), 273-278. https://doi.org/10.1080/00223980.1947.9917353

Khairat, S., Haithcoat, T., Liu, S., Zaman, T., Edson, B., Gianforcaro, R., & Shyu, C. R. (2019). Advancing health equity and access using telemedicine: A geospatial assessment. *Journal of the American Medical Informatics Association*, *26*(8-9), 796-805. https://doi.org/10.1093/jamia/ocz108

Koegel, L. K., Koegel, R. L., Ashbaugh, K., & Bradshaw, J. (2014). The importance of early identification and intervention for children with or at risk for autism spectrum disorders. *International Journal of Speech-Language Pathology, 16*(1), 50-56. https://doi.org/10.3109/17549507.2013.861511

Kornbrot, D. (2014). Point biserial correlation. *Wiley StatsRef: Statistics Reference Online*. https://doi.org/10.1002/9781118445112.stat06227

Kryszak, E. M., Albright, C. M., Fell, L. A., Butter, E. M., & Kuhlthau, K. A. (2022). Clinician perspectives on telehealth assessment of autism spectrum disorder during the COVID-19 pandemic. *Journal of Autism and Developmental Disorders, 52*, 5083-5098. https://doi.org/10.1007/s10803-022-05435-z

Landa, R. J. (2018). Efficacy of early interventions for infants and young children with, and at risk for, autism spectrum disorders. *International Review of Psychiatry, 30*(1), 25-39. https://doi.org/10.1080/09540261.2018.1432574

Lewis, L. F. (2017). A mixed methods study of barriers to formal diagnosis of autism spectrum disorder in adults. *Journal of Autism and Developmental Disorders*, *47*(8), 2410-2424. https://doi.org/10.1007/s10803-017-3168-3

Liptak, G. S., Benzoni, L. B., Mruzek, D. W., Nolan, K. W., Thingvoll, M. A., Wade, C.

M., & Fryer, G. E. (2008). Disparities in diagnosis and access to health services for children with autism: Data from the national survey of children's health. *Journal of Developmental & Behavioral Pediatrics*, *29*(3), 152-160. https://doi.org/10.1097/DBP.0b013e318165c7a0

Lockwood Estrin, G., Milner, V., Spain, D., Happé, F., & Colvert, E. (2021). Barriers to autism spectrum disorder diagnosis for young women and girls: A systematic review. *Review Journal of Autism and Developmental Disorders*, *8*(4), 454-470. https://doi.org/10.1007/s40489-020-00225-8

Lord, C., Holbrook, A., Dow, D., Byrne, K., Grzadzinski, R., Sterrett, K., Toolan, C., Kim, S. H. (2020). *Brief observation of symptoms of autism (BOSA)*. Western Psychological Services.

Lord, C., Rutter, M., DiLavore, P. C., Risi, S., Gotham, K., Bishop, S. (2012). *Autism Diagnostic Observation Schedule, second edition.* Western Psychological Services.

Lord, C., Rutter, M., & Le Couteur, A. (1994). Autism diagnostic interview-revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *Journal of Autism and Developmental Disorders, 24*(5), 659–685. https://doi.org/10.1007/BF02172145

Love, H., Panchal, N., Schlitt, J., Behr, C., & Soleimanpour, S. (2019). The use of telehealth in school-based health centers. *Global Pediatric Health*, *6*, 1-10. https://doi.org/10.1177/2333794X19884194

Ludwig, N. N., Child, A. E., Jashar, D. T., Mostow, A. J., & Wodka, E. L. (2021). Telehealth diagnosis of autism spectrum disorder through clinical cases. *The Clinical Neuropsychologist*, *36*(5), 960-980. https://doi.org/10.1080/13854046.2021.1970228

Maenner, M. J., Shaw, K. A., Bakian, A. V., Bilder, D. A., Durkin, M. S., Esler, A., Furnier, S. M., Hallas, L., Hall-Lande, J., Hudson, A., Hughes, M. M., Patrick, M., Pierce, K., Poynter, J. N., Salinas, A., Shenouda, J., Vehorn, A., Warren, Z., Constantino, J. N., ... Cogswell, M. E. (2021). Prevalence and characteristics of autism spectrum disorder among children aged 8 years—Autism and developmental disabilities monitoring network, 11 sites, United States, 2018. *MMWR Surveillance Summaries*, *70*(11), 1-16. http://dx.doi.org/10.15585/mmwr.ss7011a1

Maenner, M. J., Warren, Z., Williams, A. R., Amoakohene, E., Bakian, A. V., Bilder, D. A., Durkin, M. S., Fitzgerald, R. T., Furnier, S. M., Hughes, M. M., Ladd-Acosta, C. M., McArthur, D., Pas, E. T., Salinas, A., Vehorn, A., Williams, S., Esler, A.,

Grzybowski, A., Hall-Lande, J., ... & Shaw, K. A. (2023). Prevalence and characteristics of autism spectrum disorder among children aged 8 years—Autism and developmental disabilities monitoring network, 11 sites, United States, 2020. *MMWR Surveillance Summaries*, *72*(2), 1-14. https://doi.org/10.15585/mmwr.ss7202a1

Malik-Soni, N., Shaker, A., Luck, H., Mullin, A. E., Wiley, R. E., Lewis, M. E. S., Fuentes, J., & Frazier, T. W. (2022). Tackling healthcare access barriers for individuals with autism from diagnosis to adulthood. *Pediatric Research*, *91*, 1028-1035. https://doi.org/10.1038/s41390-021-01465-y

Mandell, D. S., Ittenbach, R. F., Levy, S. E., & Pinto-Martin, J. A. (2007). Disparities in diagnoses received prior to a diagnosis of autism spectrum disorder. *Journal of Autism and Developmental Disorders, 37*, 1795–1802. https://doi.org/10.1007/s10803-006-0314-8

Mandell, D. S., Wiggins, L. D., Carpenter, L. A., Daniels, J., DiGuiseppi, C., Durkin, M. S., Giarelli, E., Morrier, M. J., Nicholas, J. S., Pinto-Martin, J. A., Shattuck, P. T., Thomas, K. C., Yeargin-Allsopp, M., & Kirby, R. S. (2009). Racial/ethnic disparities in the identification of children with autism spectrum disorders. *American Journal of Public Health*, *99*(3), 493-498. https://doi.org/10.2105/AJPH.2007.131243

Martinez, M., Thomas, K. C., Williams, C. S., Christian, R., Crais, E., Pretzel, R., & Hooper, S. R. (2018). Family experiences with the diagnosis of autism spectrum disorder: System barriers and facilitators of efficient diagnosis. *Journal of Autism and Developmental Disorders*, *48*(7), 2368-2378. https://doi.org/10.1007/s10803-018-3493-1

Matson, J. L., Wilkins, J., & González, M. (2008). Early identification and diagnosis in autism spectrum disorders in young children and infants: How early is too early? *Research in Autism Spectrum Disorders*, *2*(1), 75-84. https://doi.org/10.1016/j.rasd.2007.03.002

Matthews, N. L., Skepnek, E., Mammen, M. A., James, J. S., Malligo, A., Lyon, A., Mitchell, M., Kiefer, S. L., & Smith, C. J. (2021). Feasibility and acceptability of a telehealth model for autism diagnostic evaluations in children, adolescents, and adults. *Autism Research*, *14*(12), 2564-2579. https://doi.org/10.1002/aur.2591

Mazurek, M. O., Handen, B. L., Wodka, E. L., Nowinski, L., Butter, E., & Engelhardt, C. R. (2014). Age at first autism spectrum disorder diagnosis: The role of birth cohort, demographic factors, and clinical features. *Journal of Developmental & Behavioral Pediatrics*, *35*(9), 561-569. https://doi.org/10.1097/DBP.0000000000000097

Meek, G. E., Ozgur, C., & Dunning, K. (2007). Comparison of the t vs. wilcoxon signed-rank test for likert scale data and small samples. *Journal of Modern Applied Statistical Methods, 6*(1), 91-106. https://doi.org/10.22237/jmasm/1177992540

Micheletti, M., Brukilacchio, B. H., Hooper-Boyle, H., Basiru, T., Brinster, M. I., Ravenscroft, S., & Shahidullah, J. D. (2023). Evaluating the efficiency and equity of autism diagnoses via telehealth during COVID-19. *Journal of Autism and Developmental Disorders*, 1-7. https://doi.org/10.1007/s10803-023-05986-9

Nazneen, N., Rozga, A., Smith, C. J., Oberleitner, R., Abowd, G. D., & Arriaga, R. I. (2015). A novel system for supporting autism diagnosis using home videos: Iterative development and evaluation of system design. *JMIR mHealth and uHealth*, *3*(2), Article e4393. https://doi.org/10.2196/mhealth.4393

NCSS Statistical Software. (n.d.). Point-Biserial and Biserial Correlations. In NCSS Statistical Software Documentation. Retrieved from ncss.com/software/ncss.

Neece, C., McIntyre, L. L., & Fenning, R. (2020). Examining the impact of COVID-19 in ethnically diverse families with young children with intellectual and developmental disabilities. *Journal of Intellectual Disability Research*, *64*(10), 739-749. https://doi.org/10.1111/jir.12769

Nickel, R. E. (2020). *The Observation of Play Screener: Home edition.* Oregon Health & Science University.

Norman, G. (2010). Likert scales, levels of measurement and the "laws" of statistics. *Advances in Health Sciences Education, 15*(5), 625-632. https://doi.org/10.1007/s10459-010-9222-y

Pierce, K., Gazestani, V. H., Bacon, E., Barnes, C. C., Cha, D., Nalabolu, S., Lopez, L., Moore, A., Pence-Stophaeros, S. & Courchesne, E. (2019). Evaluation of the diagnostic stability of the early autism spectrum disorder phenotype in the general population starting at 12 months. *JAMA Pediatrics*, *173*(6), 578-587. https://doi.org/10.1001/jamapediatrics.2019.0624

Pomales-Ramos, A., Tokish, H., Howard, M., Straiton, D., & Ingersoll, B. (2023). A mixed-methods examination of clinicians' perceived barriers to telehealth delivered applied behavior analysis. *Frontiers in Psychology*, *14*, Article 1173644. https://doi.org/10.3389/fpsyg.2023.1173644

Pompa-Craven, P., Tierman, E., Martino, J., & Lotfizadeh, A. D. (2022). Caregiver

satisfaction with delivery of telehealth autism services. *Advances in Neurodevelopmental Disorders, 6*(2), 196-205. https://doi.org/10.1007/s41252-022-00256-8

Rasch, D., Kubinger, K. D., & Moder, K. (2011). The two-sample t test: Pre-testing its assumptions does not pay off. *Statistical Papers, 52*, 219-231. https://doi.org/10.1007/s00362-009-0224-x

Reese, R. M., Jamison, R., Wendland, M., Fleming, K., Braun, M. J., Schuttler, J. O., & Turek, J. (2013). Evaluating interactive videoconferencing for assessing symptoms of autism. *Telemedicine and E-Health*, *19*(9), 671-677. https://doi.org/10.1089/tmj.2012.0312

Reese, R. J., Slone, N. C., Soares, N., & Sprang, R. (2015). Using telepsychology to provide a group parenting program: A preliminary evaluation of effectiveness. *Psychological Services*, *12*(3), 274-282. https://doi.org/10.1037/ser0000018

Reisinger, D. L., Hines, E., Raches, C., Tang, Q., James, C., & Keehn, R. M. (2022). Provider and caregiver satisfaction with telehealth evaluation of autism spectrum disorder in young children during the COVID-19 pandemic. *Journal of Autism and Developmental Disorders*, *52*, 5099-5113. https://doi.org/10.1007/s10803-022-05576-1

Rietveld, T., & van Hout, R. (2017). The paired t test and beyond: Recommendations for testing the central tendencies of two paired samples in research on speech, language and hearing pathology. *Journal of Communication Disorders*, *69*, 44-57. https://doi.org/10.1016/j.jcomdis.2017.07.002

Rosenberg, R. E., Landa, R., Law, J. K., Stuart, E. A., & Law, P. A. (2011). Factors affecting age at initial autism spectrum disorder diagnosis in a national survey. *Autism Research and Treatment*, *2011*, Article 874619. https://doi.org/10.1155/2011/874619

Rutter, M., Bailey, A., & Lord, C. (2003). *Social Communication Questionnaire (SCQ)*. Western Psychological Services.

Schober, P., Boer, C., & Schwarte, L. A. (2018). Correlation coefficients: Appropriate use and interpretation. *Anesthesia & Analgesia*, *126*(5), 1763-1768. https://doi.org/10.1213/ANE.0000000000002864

Sedgwick, P. (2012). Pearson's correlation coefficient. *British Medical Journal*, 345. https://doi.org/10.1136/bmj.e4483

Smith, C. J., Rozga, A., Matthews, N., Oberleitner, R., Nazneen, N., & Abowd, G.

(2017). Investigating the accuracy of a novel telehealth diagnostic approach for autism spectrum disorder. *Psychological Assessment*, *29*(3), 245-252. https://doi.org/10.1037/pas0000317

Sparrow, S. S., Cicchetti, D. V., & Balla, D. A. (2005). *Vineland Adaptive Behavior Scales* (2nd ed.). Circle Pines, MN: American Guidance Service.

St. Amant, H. G., Schrager, S. M., Peña-Ricardo, C., Williams, M. E., & Vanderbilt, D. L. (2018). Language barriers impact access to services for children with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *48*(2), 333-340. https://doi.org/10.1007/s10803-017-3330-y

Stahmer, A. C., Vejnoska, S., Iadarola, S., Straiton, D., Segovia, F. R., Luelmo, P., Morgan, E. H., Lee, H. S., Javed, A., Bronstein, B., Hochheimer, S., Cho, E., Aranbarri, A., Mandell, D., Hassrick, E. M., Smith, T., & Kasari, C. (2019). Caregiver voices: Cross-cultural input on improving access to autism services. *Journal of Racial and Ethnic Health Disparities, 6*, 752-773. https://doi.org/10.1007/s40615-019-00575-y

Stainbrook, J., Weitlauf, A., Ju´arez, A., Taylor, J., Hine, J., Broderick, N., Nicholson, A., & Warren, Z. (2019). Measuring the service system impact of a novel telediagnostic service program for young children with autism spectrum disorder. *Autism: The International Journal of Research and Practice, 23*(4), 1051–1056. https://doi.org/10.1177/ 1362361318787797

Stavropoulos, K.K.M., Bolourian, Y., & Blacher, J. (2022) A scoping review of telehealth diagnosis of autism spectrum disorder. *PLoS ONE*, *17*(2), Article e0263062. https://doi.org/10.1371/journal.pone.0263062

Stavropoulos, K. M., Heyman, M., Salinas, G., Baker, E., & Baker, J. (2021). Exploring telehealth during COVID for assessing autism spectrum disorder in a diverse sample. *Psychology in the Schools*, *59*(7), 1319-1334. https://doi.org/10.1002/pits.22672

Sutherland, R., Trembath, D., & Roberts, J. (2018). Telehealth and autism: A systematic search and review of the literature. *International Journal of Speech-Language Pathology*, *20*(3), 324-336. https://doi.org/10.1080/17549507.2018.1465123

Tai, D. B. G., Shah, A., Doubeni, C.A., Sia, I. G., & Wieland, M. L. (2020). The disproportionate impact of COVID-19 on racial and ethnic minorities in the United States. *Clinical Infectious Diseases, 72*(4), 703-706. https://doi.org/10.1093/cid/ciaa815

Talbott, M. R., Lang, E., Avila, F., Dufek, S., & Young, G. (2022). Experiences of

caregivers participating in a telehealth evaluation of development for infants (TEDI). *Journal of Autism and Developmental Disorders, 52*, 5266-5273. https://doi.org/10.1007/s10803-022-05607-x

Turner, L. M., & Stone, W. L. (2007). Variability in outcome for children with an ASD diagnosis at age 2. *Journal of Child Psychology and Psychiatry*, *48*(8), 793-802. https://doi.org/10.1111/j.1469-7610.2007.01744.x

Wagner, L., Corona, L. L., Weitlauf, A. S., Marsh, K. L., Berman, A. F., Broderick, N. A., Francis, S., Hine, J., Nicholson, A., Stone, C. & Warren, Z. (2021a). Use of the TELE-ASD-PEDS for autism evaluations in response to COVID-19: Preliminary outcomes and clinician acceptability. *Journal of Autism and Developmental Disorders*, *51*(9), 3063-3072. https://doi.org/10.1007/s10803-020-04767-y

Wagner, L., Stone, C., Wade, J., Corona, L., Hine, J., Nicholson, A., Swanson, A., Vehom, A., Weitlauf, A., & Warren, Z. (2021b). *TELE-ASD-PEDS: A Telemedicine-Based ASD Evaluation Tool for Toddlers and Young Children: TELE-ASD-PEDS User's Manual.* Vanderbilt University Medical Center.

Wagner, L., Weitlauf, A. S., Hine, J., Corona, L. L., Berman, A. F., Nicholson, A., Allen, W., Black, M. & Warren, Z. (2022). Transitioning to telemedicine during COVID-19: Impact on perceptions and use of telemedicine procedures for the diagnosis of autism in toddlers. *Journal of Autism and Developmental Disorders*, *52*, 2247-2257. https://doi.org/10.1007/s10803-021-05112-7

Walfish, S. (2006). A review of statistical outlier methods. *Pharmaceutical Technology, 30*(11), 82-88.

Warren, Z., McPheeters, M. L., Sathe, N., Foss-Feig, J. H., Glasser, A., & Veenstra-VanderWeele, J. (2011). A systematic review of early intensive intervention for autism spectrum disorders. *Pediatrics*, *127*(5), e1303-e1311. https://doi.org/10.1542/peds.2011-0426

Wechsler, D. (2008). *Wechsler Adult Intelligence Scale—fourth edition. Technical and interpretive manual*. Psychological Corporation.

Wechsler, D. (2011). *Wechsler Abbreviated Scale of Intelligence–Second Edition (WASI-II).* NCS Pearson.

Wiggins, L. D., Bakeman, R., Adamson, L. B., & Robins, D. L. (2007). The utility of the social communication questionnaire in screening for autism in children referred for early intervention. *Focus on Autism and Other Developmental Disabilities*, *22*(1), 33-38. https://doi.org/10.1177%2F10883576070220010401

Zuckerman, K. E., Lindly, O. J., Reyes, N. M., Chavez, A. E., Macias, K., Smith, K. N., & Reynolds, A. (2017). Disparities in diagnosis and treatment of autism in Latino and non-Latino white families. *Pediatrics*, *139*(5). https://doi.org/10.1542/peds.2016-3010