Title

Assessment of Autism Spectrum Disorder.

Journal
Assessment, 31(1)

Authors
Yu, Yue
Ozonoff, Sally
Miller, Meghan

Publication Date
2024

DOI
10.1177/10731911231173089

Peer reviewed

Assessment
2024, Vol. 31(1) 24–41

DOI: 10.1177/10731911231173089
journals.sagepub.com/home/asm
Sage

Assessment of Autism Spectrum Disorder

Yue Yu[1], Sally Ozonoff[1], and Meghan Miller[1]

Abstract

Autism Spectrum Disorder (ASD) is a neurodevelopmental condition characterized by challenges in social interaction and communication and the presence of restricted interests and repetitive behaviors. The importance of early detection of ASD and subsequent early intervention is well documented. Efforts have been made over the years to clarify ASD diagnostic criteria and develop predictive, accurate screening tools and evidence-based, standardized diagnostic instruments to aid in the identification of ASD. In this article, we review the most recent changes in ASD diagnostic criteria in the *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition, Text Revision*, summarize evidence-based instruments for ASD screening and diagnostic evaluations as well as the assessment of co-occurring conditions in ASD, the impact of COVID-19 on ASD assessment, and directions for future research in the field of ASD assessment.

Keywords

autism, *DSM-5-TR*, early screening, diagnostic assessment, COVID-19

Autism Spectrum Disorder (ASD) is a neurodevelopmental condition characterized by challenges in social interaction and communication, and the presence of restricted interests and repetitive behaviors (*Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition, Text Revision* [*DSM-5-TR*]; American Psychiatric Association [APA], 2022). Recent data indicate that one in 44 children in the United States (Maenner et al., 2021) and one in 160 children internationally (World Health Organization, 2013) has a diagnosis of ASD. Critical to accurate diagnosis and appropriate treatment planning is an evidence-based approach to assessment. Consistent with this special issue's focus on evidence-based assessment guidelines for both clinical practice and research, this article provides a synthesis of available literature that describes the range of evidence-based evaluation approaches for ASD, from screening to diagnosis. We begin with reviewing the most commonly used, evidence-based instruments for screening and diagnostic evaluation of ASD. We then discuss the importance and implications of early screening and diagnosis, consider best practices for the evaluation of co-occurring conditions, and describe the impact of COVID-19 on recent autism diagnostic practices. We end the article with future directions for novel approaches to autism assessment.

At the outset, we wish to highlight recent discussions in the autism community related to identity-first versus person-first language (Botha et al., 2021; Vivanti, 2020). Caregivers of autistic people and professionals tend to prefer person-first language (i.e., individuals with autism) to indicate that the disorder does not define the person (Buijsman et al., 2022; Robison, 2019). Autistic advocates, on the contrary, have recently emphasized that the use of person-first language may perpetuate stigma, dehumanization, and violence (Botha et al., 2021). In the autism community, there currently is no clear consensus on the most preferred terminology (Botha et al., 2021), but most self-advocates prefer identity-first language (Gernsbacher, 2017). We also recognize the variability in terminology preferences across stakeholders. As such, in this article, we use ASD when referring to the diagnostic characterization used by service systems. We use person-first and identity-first terminology interchangeably to reflect both the preferences of autistic individuals and self-advocates and of many families of young children.

Changes in Diagnostic Criteria Over Time

According to the *DSM-5* (APA, 2013), to meet the diagnostic criteria, an individual must exhibit persistent

[1]University of California, Davis, California, USA

Corresponding Author:
Yue Yu, MIND Institute, University of California, Davis, 2825 50th Street, Sacramento, CA 95817, USA.
Email: eeyu@ucdavis.edu

Table 1. Diagnostic Criteria of Autism Spectrum Disorder in the DSM-5.

A. Social communication and interaction challenges:
1. Deficits in social-emotional reciprocity
2. Deficits in nonverbal communicative behaviors used for social interaction
3. Deficits in developing, maintaining, and understanding relationships
B. Repetitive or stereotyped behaviors:
1. Stereotyped or repetitive motor movements, use of objects, or speech
2. Insistence on sameness, inflexible adherence to routines, or ritualized patterns of verbal or nonverbal behavior
3. Highly restricted, fixated interests that are abnormal in intensity or focus
4. Hyper- or hypo-reactivity to sensory input or unusual interest in sensory aspects of the environment

Note. DSM-5 = *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition.*

challenges in social communication and interaction across multiple settings and also demonstrate symptoms in the restricted, repetitive patterns of behaviors, interests, or activities domain (see Table 1).

The diagnostic criteria for autism have evolved in the past several decades and across the previous editions of the *DSM*. The most recent major changes in autism diagnosis were in the transition from *DSM* (4th ed., text rev.; *DSM-IV-TR*; APA, 2000) to *DSM-5*. In this revision process, previously separate diagnoses (i.e., autistic disorder, Asperger's disorder, pervasive developmental disorder not otherwise specified) were consolidated into one: ASD. In addition, the three symptom categories of social impairment, language/communication impairment, and repetitive/restricted behaviors were merged into two symptom domains: persistent deficits in social communication and interaction and restricted, repetitive patterns of behaviors. This change was meant to reflect the state of the science which consistently revealed that the social and communication/language symptoms in *DSM-IV* (APA, 1994) loaded on a single "social communication" factor (Mandy et al., 2012). Two other key changes involved the addition of sensory symptoms (both hypo- and hyper-reactivity) within the category of restricted and repetitive behaviors and the inclusion of a severity assessment scale (Levels 1–3) based on the level of support needed for daily functioning in autistic individuals.

Changes from the *DSM-5* to *DSM-5-TR* have been much more minimal, although the two small changes are worth noting. Since the *DSM-5*'s initial release in 2013, there has been some ambiguity around the diagnostic criteria in the social communication domain. Specifically, while it was clear that at least two of the four restricted and repetitive behavior symptoms were required to meet the diagnostic criteria, the minimum required symptoms in the social communication domain were somewhat ambiguous (i.e., whether *any* or *all* three symptoms are required; Hess, 2022). The recently released *DSM-5-TR* (APA, 2022) made modifications to the diagnostic criteria for ASD to clarify this, now specifying that *all three* social communication symptoms should be present. The second minor modification involves a change in the wording of one of the specifiers. Specifically, "another neurodevelopmental, mental, or behavioral **disorder**" is now "another neurodevelopmental, mental, or behavioral **problem**." Thus, this specifier has been broadened such that it no longer needs to be a diagnosable condition.

Because of the increased awareness of autism in both the general population and the scientific community, along with broadened diagnostic criteria, some individuals are being diagnosed for the first time in adulthood; this group is sometimes described as the "lost generation" (Lai & Baron-Cohen, 2015). These adults may have been undiagnosed or misdiagnosed when they were young. Initial autism identification in adulthood is challenging for various reasons, such as a lack of or inaccurate history of early development, skills developed to mask social difficulties, and other potential co-occurring, possibly more debilitating, psychiatric comorbidities (Lai & Baron-Cohen, 2015). Diagnosing autism in late life is even more challenging, for similar reasons (Van Niekerk et al., 2011). Although the awareness of autism identification in adulthood has improved over the years, and there has been an increase in autism research focused on older age populations (Mason et al., 2022), supporting individuals on the spectrum throughout their lifespan requires ongoing effort (Mason et al., 2022; Piven & Rabins, 2011). Individuals diagnosed in adulthood are more likely to report having psychiatric conditions compared with individuals diagnosed in childhood (Jadav & Bal, 2022), which emphasizes the need for valid and reliable methods for early diagnosis.

Assessment and Early Identification of Autism

The ability to accurately diagnose autism has improved considerably in the past two decades, especially in very young children. This has been the direct result of a number of empirical investigations focused on the development and/or adaptation of reliable and valid assessment instruments (Zwaigenbaum & Penner, 2018). A critical implication of early detection is the ability to refer young children to intervention services as early as possible. Indeed, the importance of early detection of autism and subsequent early intervention is well documented (Elder et al., 2017; Hyman et al., 2020; Rogers et al., 2012; Rotholz et al., 2017). At the same time, valid concerns exist around false positives among children who are diagnosed at very young ages. Thus, it is important to investigate the stability of the diagnosis (i.e., the likelihood that a child will still meet diagnostic criteria at follow-up; Charman & Baird, 2002) and developmental trajectories of symptoms, which may have implications both scientifically and clinically (Ozonoff et al., 2015; Woolfenden et al., 2012).

Research has shown that symptoms of autism typically begin to emerge between 12 and 18 months of age (Landa et al., 2013; Ozonoff et al., 2010; Pierce et al., 2019) and that diagnoses made at 18 months are reliable and stable (Ozonoff et al., 2015). Numerous studies have demonstrated the stability of diagnoses among children diagnosed before three (Chawarska et al., 2009; Guthrie et al., 2013; Ozonoff et al., 2015), suggesting that the false-positive rate of diagnosis before age 3 is relatively low. The false-negative rate is higher, however. In a longitudinal study in which serial diagnostic assessments were conducted, close to half of children diagnosed with autism at age 3 did not meet diagnostic criteria at 24 months of age (Ozonoff et al., 2015). And a small group of children appear to have even later emerging symptoms (Davidovitch et al., 2015; Michelotti et al., 2002; Ozonoff et al., 2018); these children did not meet the criteria by the age of 3 but did meet the criteria at school-age (late diagnosed). These late-diagnosed cases are generally heterogeneous in their early outcome classifications (e.g., typically developing, subclinical symptoms) and phenotype (Ozonoff et al., 2018). It is speculated that diagnostic overshadowing by other conditions (early delays in language or cognitive function; Davidovitch et al., 2015) and/or a prolonged period of symptom development (Ozonoff et al., 2018) may contribute to late diagnosis. These findings suggest the need to expand autism screening and/or surveillance beyond three years of age. Below, we review screening and diagnostic instruments that can be utilized across these developmental periods.

Empirically Validated Screening and Diagnostic Measures

When selecting assessment instruments, it is important to consider their empirical evidence, standardization across large populations, and psychometric properties, such as reliability, validity, sensitivity (i.e., the proportion of autistic individuals correctly identified by the screening instrument as having autism), specificity (i.e., the proportion of non-autistic individuals correctly identified as not having autism), and positive-predictive value (i.e., the proportion of individuals over the screening cutoff who are ultimately diagnosed with autism). In the following section, we first review commonly used, evidence-based autism screening instruments followed by an overview of autism-specific diagnostic tools, developmental and intellectual assessments, and evaluation of adaptive behavior.

Autism Screening

Screening for autism involves the use of brief assessments to determine the level of likelihood of a diagnosis. The American Academy of Pediatrics recommends that all children be screened for autism at 18 and 24 months (Hyman et al., 2020) which can be accomplished through standardized caregiver reports or direct observational instruments. While screening and surveillance are both used to detect early developmental concerns, these approaches differ. Developmental surveillance refers to continuous monitoring and collection of information that may relate to certain developmental concerns over time (e.g., having a sibling with ASD, parental concern, pediatrician concern), which may or may not involve standardized instruments, whereas autism screening is accomplished through the administration of specific standardized, evidence-based tools (e.g., M-CHAT; Council on Children with Disabilities). If a child is determined to be at risk of autism, referrals should be made for a more in-depth diagnostic evaluation. In cases that are clear (whether in terms of high screening scores or other developmental delays), children may be directly referred to early intervention services prior to a comprehensive evaluation.

Efforts have been made over the years to develop predictive, accurate measures to best identify young children at risk of autism as early as possible. These tools are generally categorized as *Level 1* or *Level 2* screeners (Petrocchi et al., 2020). *Level 1* screeners are brief instruments used to screen children at risk of autism in the general population and are usually completed by parents and/or administered by pediatricians during well-child visits, whereas *Level 2* screeners are used to screen children who are already identified as having elevated likelihood for autism (e.g., children with previous developmental concerns, younger siblings of children with autism; Hardy et al., 2013; Petrocchi et al., 2020). Below, we describe several commonly used screening tools.

Modified Checklist for Autism in Toddlers. The Modified Checklist for Autism in Toddlers (M-CHAT; Robins et al., 2014) is one of the most widely used *Level 1* autism screening instruments. The revised and updated version, M-CHAT, Revised with Follow-Up (M-CHAT-R/F; Robins et al., 2014) is a two-stage screener designed for toddlers ages 16 to 30 months. First, caregivers complete a series of 20 yes/no questions that assess early social communication skills and other signs of autism. If the total score is ≤ 2, no follow-up is needed (low risk). A score between 3 and 7 suggests moderate risk and a structured follow-up interview is administered to caregivers to obtain additional information. If the total score of the follow-up is 2 or higher, the child should be referred for diagnostic evaluation and early intervention. A score of 8 or higher indicates high risk, and the follow-up interview can be bypassed in favor of immediate referral for early intervention services and diagnostic evaluation. The M-CHAT-R/F takes about 10 min to complete.

The validation sample for the M-CHAT-R/F was large (Robins et al., 2014), and the internal consistency reliability of M-CHAT-R/F is acceptable (Cronbach's $\alpha = 0.79$; Robins et al., 2014) as are sensitivity (0.85) and specificity (0.99) values (Robins et al., 2014). The positive-predictive value (PPV) is 0.48 for autism and 0.95 for any developmental disorder (Robins et al., 2014), indicating that 48% of the children who screened positive received an ASD diagnosis and 95% of the children had some sort of developmental delay or concerns, including ASD, that warranted referrals to early intervention services (Robins et al., 2014). Similar rates have been found in other countries (PPV for ASD = 0.55, PPV for any developmental disorder = 0.91; Guo et al., 2019). This indicates that the predictive value of the M-CHAT-R/F for any developmental concerns is higher than for autism specifically. Nevertheless, it should be noted that PPVs are impacted by the base rate of the disorder in the sample studied, such that when the prevalence of a disorder is higher, the probability of having the disorder after a positive test is higher (Monaghan et al., 2021). Robins et al. (2014) do not recommend using M-CHAT-R/F as a broader screening tool because its sensitivity for non-ASD delays has not been well studied (e.g., sensitivity for global developmental delays, communication delays, etc.). The M-CHAT-R/F has been translated into more than 40 languages (www.mchatscreen.com) and has demonstrated satisfactory sensitivity and specificity in adaptations in some countries and regions (e.g., Spain, China, Taiwan; Dai et al., 2021). Screening accuracy was shown to be similar across race and ethnic groups in a recent large study (Guthrie et al., 2019).

Social Communication Questionnaire. The Social Communication Questionnaire (SCQ; Rutter et al., 2003) is a 40-item yes/no rating scale completed by caregivers. It is a *Level 2* screener used to evaluate reciprocal social interaction, language and communication, and stereotyped patterns of behavior in children older than 4 years of age with a mental age of more than 2 years. The SCQ has two forms: Lifetime (i.e., assess developmental history) and Current (i.e., assess behaviors in the last three months).

This measure has good overall accuracy in identifying autism with a score of 15 as the cutoff (sensitivity = 0.85, specificity = 0.75; Norris & Lecavalier, 2010), good internal consistency reliability (Cronbach's $\alpha = 0.87$; Rutter et al., 2003), good discriminative validity (0.86) in children with autism vs. non-autism diagnosis (Berument et al., 1999), and adequate concurrent validity with gold-standard diagnostic interviews for autism (0.78; Berument et al., 1999). It is available in more than 10 languages, and its psychometrics have also been examined in other regions (e.g., Gau et al., 2011; Karaminis & Stavrakaki, 2022).

Infant Toddler Checklist. The Infant Toddler Checklist (ITC; Wetherby et al., 2008) is a caregiver report measure designed to screen for communication delays in young children between 6 and 24 months of age. It consists of 24 items about social communication development. The measure yields three composite scores (social, speech, symbolic) and a total score. A score in the bottom 10^{th} percentile on the social, symbolic, or total composites indicates a positive screen for communication delays, including developmental delay and ASD. It has also been used and evaluated as a *Level 1* screener for ASD (Pierce et al., 2019). The ITC takes about 5 min to complete by caregivers.

The ITC's test-retest reliability is good (0.84 for total score; Wetherby et al., 2002). The ITC has moderate-to-strong correlations with face-to-face child evaluations (Communication and Symbolic Behavior Scales—Developmental Profile [CSBS-DP] Behavioral Sample; 0.67 at 18–24 months) and more in-depth caregiver interviews (CSBS-DP Caregiver Questionnaire; 0.87 at 18–24 months; Wetherby et al., 2002). The ITC's psychometric properties for autism screening are low in the first year of life (sensitivity = 0.20–0.64, specificity = 0.42–0.60; Parikh et al., 2021; Wetherby et al., 2008) but improve considerably by 21 to 24 months (sensitivity = 0.78–0.95, specificity = 0.83; Parikh et al., 2021; Wetherby et al., 2008).

Screening Tool for Autism in Toddlers. The Screening Tool for Autism in Toddlers (STAT; Stone et al., 2004) is a *Level 2* screening measure administered by a trained professional. It comprises 12 play-based activities to observe early social communication behaviors in children from 24 to 35 months of age that can be administered in 15 to 20 min. The total score is derived by summing the four domain scores (play, requesting, directing attention, and motor imitation), with higher scores representing greater impairment. Sensitivity for autism classification is estimated to be 0.92, and specificity is 0.85, with a test–retest reliability of 0.90 (Stone et al., 2004). The STAT has shown good discriminant validity between autism and non-spectrum (0.95; Stone et al., 2004). Nevertheless, the original study sample size was relatively small. Recent research examining its psychometrics is limited.

Emerging Measures. Several *Level 2* screening instruments that have not yet been clinically validated are available for use in research studies, particularly for infants. For example, the Autism Observation Scale for Infants (AOSI; Bryson et al., 2008) is a 19-item, direct observational measure designed to detect and monitor signs of autism in infants 6 to 18 months of age. It is a play-based instrument with semi-structured activities administered by experienced examiners and takes 10-15 minutes to administer. The items are rated either on a 0 to 3 scale (0 = *typical*, 1–3 = *increasing severity of impairment*) or a dichotomized scale (0 = *typical*, 2 = *atypical*). A total score and a total number of items endorsed (i.e., marker counts) are calculated. Initially developed for research purposes, it has been most often used in high-risk sibling studies. According to a recent study (Zwaigenbaum, Bryson, et al., 2021), the AOSI's sensitivity is low at 12 (0.52, with 0.74 specificity, cutoff = 7) and 15 months (0.41, with 0.90 specificity, cutoff = 10), better at 6 (0.57, with 0.51 specificity, cutoff = 7), 9 months (0.60, with 0.53 specificity, cutoff = 6), and best at 18 months (0.73, with 0.65 specificity, cutoff = 6).

Approaches to Autism Evaluation

When screening results suggest risk of autism, children are typically referred for more comprehensive, formal diagnostic evaluations. Autism diagnostic assessment practices have made considerable advancements over the past several decades. Moving away from relying solely on clinical judgment, the field has focused on developing and evaluating evidence-based, standardized assessment instruments to better inform clinical judgment and identify autism (Zwaigenbaum, Bishop, et al., 2021).

Individuals referred for autism evaluations often experience challenges in other areas beyond the core autism symptom domains, including developmental or cognitive delays, language delays, co-occurring medical and mental health difficulties, and behavioral concerns. Consequently, it is important to consider various developmental and psychiatric domains that may impact symptom manifestations and reveal individual needs in addition to diagnostic classification (Steiner et al., 2012; Whitehouse et al., 2018; Zwaigenbaum & Penner, 2018). As such, it is often recommended to include an interdisciplinary team of professionals (e.g., psychologist, speech-language pathologist, and developmental–behavioral pediatrician). However, although it could be comprehensive and convenient for families, a multi/interdisciplinary evaluation approach is not always covered by third-party payors, not available in all health systems or geographic locations, and not always efficient for uncomplicated cases (Penner et al., 2018). For example, the Australian national guidelines for the assessment and diagnosis of autism recommend a multidisciplinary team approach specifically for more complex or ambiguous cases (Whitehouse et al., 2018).

A comprehensive evaluation should include informant report from various sources (e.g., caregiver, teacher, self-report) and direct observation by a trained provider. It often begins with a structured or semi-structured caregiver interview, through which providers gather general and autism-specific information, including developmental milestones, medical and behavioral history, caregiver current concerns, social communication challenges, restricted and repetitive behaviors and interests, and mental health and behavioral concerns. In addition to the evaluations and observations that target autism-related symptoms, other broad developmental areas (e.g., cognitive and adaptive abilities) need to be examined as well to make differential diagnoses or assess for potential co-occurring conditions. Primary goals of a comprehensive autism evaluation typically involve making an accurate diagnosis (Risi et al., 2006), identifying co-occurring conditions, profiling individual strengths and challenges, clarifying personal or family interests and goals, making service and treatment recommendations, and identifying areas that need further evaluation (e.g., vision, hearing, genetic testing; Powell et al., 2018). In the following section, we review commonly used, empirically supported caregiver interviews, direct observational instruments, cognitive/developmental assessment tools, and adaptive measures.

Caregiver Interview

Autism Diagnostic Interview-Revised. The Autism Diagnostic Interview-Revised (ADI-R; Lord et al., 1994) is a comprehensive, semi-structured caregiver interview often used as part of a comprehensive evaluation to inform clinical judgment with respect to an autism diagnosis, in combination with other measures. The ADI-R was originally developed based on the *DSM-IV-TR* and the International Classification of Diseases-10 diagnostic criteria. It assesses current and historical behaviors related to autism (from 18 months to adulthood) in individuals with a mental age of at least 2 years. Three algorithm scores are generated: social difficulty, communication challenges, and repetitive behaviors. Administration time ranges from 90 to 150 min by a trained clinician. A toddler version of the ADI-R designed for children under 4 years of age includes 32 new questions and a new algorithm (S. H. Kim et al., 2013); sensitivity (0.89–0.97) and specificity (0.86–0.89) for this algorithm are strong.

The ADI-R has good interrater reliability (Zander et al., 2017). The overall sensitivity and specificity for the ADI-R are estimated to be 0.75 and 0.82, respectively, but estimates of sensitivity (0.33–1.00) and specificity (0.61–1.00) vary widely among studies, according to a recent meta-analysis (Lebersfeld et al., 2021). Lebersfeld et al. (2021) also found that studies using combined research and clinical samples (0.82) had higher sensitivity than clinical-only (0.71) or research-only samples (0.73). Research samples (0.85) had the highest specificity, followed by combined samples (0.76) and clinical samples (0.72). Thus, it seems that the ADI-R is most accurate when used in research studies. Due to its length and training requirements, the ADI-R is mostly used for research. In practice, many providers rely on (non-standardized) interviews to collect information relevant to each *DSM* criterion.

Behavioral Observation

In addition to caregiver interviews, observational measures are an important part of a comprehensive diagnostic evaluation process. The Autism Diagnostic Observation Schedule (ADOS; Lord et al., 1999, 2012) is considered the "gold-standard" semi-structured, interaction-based, observational evaluation for autism. The Pre-linguistic ADOS (PL-ADOS; DiLavore et al., 1995) was developed to assess younger children (younger than 6 years of age) and/or preverbal children. The ADOS-Generic (ADOS-G; Gotham et al., 2007; Lord et al., 1999), with revised algorithms, led to increased diagnostic validity and reliability and was designed to assess communication, social interaction, and repetitive behaviors and interests. The Autism Diagnostic Observation Schedule-2nd Edition (ADOS-2; Lord et al., 2012), the most recent version, revised algorithms based on the ADOS-G (Gotham et al., 2007), introduced a new comparison score for Modules 1 through 3 and added a Toddler Module. The ADOS-2 consists of five modules, the selection of which depends on the individual's expressive language and developmental level. It can be used with individuals who have a nonverbal developmental age of at least 12 months through adulthood and takes approximately 40 to 60 min to administer. The Toddler Module (12 to 30 months) and Module 1 (31 months and older) are for children who do not use phrase speech consistently. Module 2 is for children who use phrase speech but are not yet verbally fluent, regardless of age. Modules 3 (children and adolescents for whom playing with action figures are appropriate) and 4 (older adolescents and adults) are designed for individuals with fluent speech. Diagnostic algorithms yield two domain scores: social affect and restricted and repetitive behaviors, along with a total score. Modules 1 to 4 have three diagnostic classifications: non-spectrum, ASD, and Autism. The ADOS-2 Toddler module yields three ranges of concerns: little-to-no concern, mild-to-moderate concern, and moderate-to-severe concern.

The ADOS-2 has been shown to have good interrater reliability: 0.79 to 0.98 across the five modules (Lord et al., 2012). A recent meta-analysis (Lebersfeld et al., 2021) found that the overall sensitivity and specificity of the ADOS-2 and its direct precursors (ADOS-Toddler; Luyster et al., 2009, and ADOS-G with revised algorithms) ranged from 0.89 to 0.92 and 0.81 to 0.85, respectively. Similar to the ADI-R, research samples had higher levels of accuracy than clinical and combined samples (Lebersfeld et al., 2021). Consistent with the recommendation of performing comprehensive evaluations, studies have shown that the diagnostic accuracy is higher when both the ADOS and the ADI-R are used together (sensitivity 0.70–0.98, specificity 0.80–0.95) than when they are used individually (S. H. Kim & Lord, 2012; Risi et al., 2006).

Childhood Autism Rating Scale-2 (CARS-2). The *Childhood Autism Rating Scale-2* (CARS-2; Schopler et al., 2010) is a clinician rating scale that involves 15 items, each rated on a four-point scale (1 = *within normal limits* to 4 = *severely abnormal for age*) based on observation in the areas of social, emotional, adaptive, communication, and cognitive functioning. It consists of a Standard Form (CARS-2-ST), meant for use in children ages 2 - 6 with an estimated IQ of 79 and lower, and a "High Functioning" form (CARS-2-HF), meant for use in those ages 6 and above with an estimated IQ of 80 or higher. A total raw score is classified into three categories: minimal to no symptoms of ASD, mild to moderate, and severe symptoms.

The CARS-2-ST and CARS-2-HF have been shown to have acceptable sensitivity (0.84, 1.00) and specificity

(1.00, 0.71; Dawkins et al., 2016), good internal consistency reliability (Cronbach's α = 0.90; Moon et al., 2019), and good correspondence with *DSM-5* classification (Moulton et al., 2019), despite being developed originally based on *DSM* (3rd ed.; *DSM-III*; APA, 1980) conceptualizations of ASD. The CARS-2 is published in seven other languages and takes about 5 to 10 min to complete after an observation is conducted by a provider.

Developmental and Intellectual Functioning Measures

An individual's cognitive or developmental functioning provides an important context in which to understand their social communication and behavior. Therefore, it is important to assess developmental functioning as part of any comprehensive evaluation for autism. Here, we focus only on directly administered standardized assessments. When selecting an instrument, numerous factors must be considered, such as the appropriateness of the measure to both the individual's chronological and mental age, the standardization sample, and the ability to separately evaluate verbal and nonverbal skills. The purpose of the assessment, clinicians' skills, and family's goals should also help guide which instruments to use.

It is worth noting that developmental assessments and intellectual assessments are not synonymous. Developmental assessments examine early learning and cognitive development in young children (i.e., infants and toddlers), including visual–spatial, motor, and early language and communication functioning. Results of developmental assessment can inform the need for further evaluation or early intervention, but the resulting summary scores are not equivalent to intelligence quotient (IQ) scores as they measure different skills. Moreover, intellectual functioning is less stable earlier in development, with the potential for large changes over time (Johnson & Marlow, 2006). The target population for intellectual testing, on the contrary, is older (preschool-aged and above). Intellectual testing is used to assess global intelligence, including nonverbal, processing speed, and working memory abilities, and provides IQ scores.

Commonly used, evidence-based developmental or cognitive assessments include Bayley Scales of Infant Development—Fourth edition (Bayley & Aylward, 2019), Mullen Scales of Early Learning (MSEL; Mullen, 1995), and Differential Abilities Scale—Second Edition (DAS-II; Elliot, 2007). Commonly used IQ tests include the Wechsler Intelligence Scales (Wechsler, 2011, 2012, 2014), Stanford-Binet (SB; Roid, 2003), and, for nonverbal individuals, the Leiter International Performance Scale (Leiter; Roid et al., 2013).

Developmental/Cognitive Assessment

Bayley Scales of Infant Development. The Bayley is a widely used developmental assessment instrument that takes approximately 30 to 70 min to complete (Bayley & Aylward, 2019). It examines cognitive, adaptive, social–emotional, fine and gross motor, and expressive and receptive language functioning in children aged 16 days to 42 months and is used to identify developmental delays. The Bayley-4 has strong internal consistency (0.85–0.99; Bayley & Aylward, 2019) and good test–retest reliability across the composite scores (0.81–0.85; Bayley & Aylward, 2019). Correlations between the Bayley-4 and Wechsler Preschool and Primary Scale of Intelligence, Fourth edition (WPPSI-IV; Wechsler, 2012) range from 0.64 – 0.79 across scales (Bayley & Aylward, 2019). A strength of this instrument is that it has updated norms and a large standardization sample, including individuals with autism (Bayley & Aylward, 2019).

Mullen Scales of Early Learning. The MSEL (Mullen, 1995) is another commonly used standardized developmental assessment for children birth to 68 months of age. It takes approximately 20 to 60 minutes to complete and has been widely used in research studies of emerging ASD and the early ASD phenotype (e.g., Bishop et al., 2011; Werner & Dawson, 2005). Subscales include gross and fine motor, expressive and receptive language, and visual reception skills. The MSEL has several strengths. First, it has good internal, test–retest, and interrater reliability (Mullen, 1995) and good construct validity, convergent validity, and divergent validity (Swineford et al., 2015). It also allows for separate scores to be generated in verbal and nonverbal domains as well as distinguishes receptive and expressive language. In addition, the wide age range allows for the tracking of developmental progress in children over time. Nevertheless, autistic children were not included in the standardization sample, and the norms of the measure have not been updated in nearly 30 years.

Differential Abilities Scale—Second Edition. The DAS-II (Elliot, 2007) is designed to assess cognitive abilities in youth aged 2 years, 6 months to 17 years, and 11 months. It comprises two overlapping batteries, the Early Years (2:6–8:11) and the School Age battery (5:0–17:11). It is designed to measure verbal reasoning, nonverbal reasoning, and spatial abilities. This instrument generates a General Conceptual Ability (score and cluster scores. For children older than three and half, a Special Nonverbal Composite score can be generated for nonverbal core subtests. This can be useful in assessing cognitive abilities in non-speaking children with autism (Ozonoff et al., 2005; Sutera et al., 2007). Instructions are available in Spanish (Elliot, 2007, 2012). The DAS-II can be used to identify personal strengths and challenges and takes 45 to 60 mins to complete. Average internal consistency reliability ranges from 0.68 - 0.97, and the test–retest reliability ranges from 0.63 to 0.92. Studies have indicated good convergent validity between the MSEL and the DAS-II (Bishop et al., 2011; Farmer et al., 2016). Although individuals with ASD were not included in the standardization sample, studies have explored its utility in this population (e.g., Kuriakose, 2014).

Intellectual Assessment

Wechsler Intelligence Scales. The Wechsler Intelligence Scales are some of the most widely used IQ tests, including the WPPSI-IV (Wechsler, 2011, 2012, 2014), the Wechsler Intelligence Scale for Children—Fifth Edition (WISC-V; Wechsler, 2014), and the Wechsler Abbreviated Scale of Intelligence—Second edition (WASI-II; Wechsler, 2011). They differ by age range and subtests: The age range for WPPSI is 2 years, 6 months to 7 years, 7 months, and takes 30 to 60 min to complete, while the age range for WISC is 6 years to 16 years, 11 months and takes about 60 min to complete. Both tests include scales measuring verbal comprehension, visual–spatial skills, working memory, fluid reasoning, and processing speed. These tests also provide a comprehensive IQ score, in addition to the subscale scores, and individuals with ASD were included in the standardized samples. The WASI-II is an abbreviated intelligence measure for individuals ages 6 years to 90 years, and 11 months. The WASI-II includes only the verbal comprehension and perceptual reasoning composites and provides an *estimated* IQ score (either 2-subtest or 4-subtest) rather than a comprehensive IQ score. It takes about 30 min to complete and can be used in clinical and research settings. Individuals with ASD were not included in the standardization sample. The Weschler Intelligence tests have been shown to have acceptable to good internal consistency reliability (WPPSI-IV: 0.86–0.96; WISC-V: 0.77–0.93; WASI-II: 0.87–0.97), test–retest reliability (WPPSI-IV: 0.75–0.87; WISC-V: 0.80–0.91; WASI-II: 0.79–0.96), interrater reliability (WPPSI-IV: 0.98–0.99; WISC-V: 0.97–0.99; WASI-II: 0.94–0.99), and adequate to strong correlations with other measures of intelligence (WPPSI-IV: 0.65–0.86 with WPPSI-III; WISC-V: 0.74–0.81 with WISC-IV and WPPSI-IV; WASI-II: 0.71–0.92 with WASI, WISC-IV, and WAIS-IV; Wechsler, 2011, 2012, 2014).

Stanford-Binet - 5th Edition (SB-5). The *Stanford-Binet - 5th Edition* (SB-5; Roid, 2003) is used to assess intellectual abilities in individuals ages 2 to 85. It generates five-factor scores in Fluid Reasoning, Knowledge, Quantitative Reasoning, Visual-Spatial Process, and Working Memory. The SB-5 provides both verbal and nonverbal IQ scores, which are especially useful in testing autistic individuals who are non-speaking or have limited language abilities. It takes 45 to 75 min to complete. Individuals with ASD were included in the standardization sample. The SB-5 has good internal consistency reliability (0.98; DiStefano & Dombrowski, 2006), test–retest reliability (0.86; DiStefano & Dombrowski, 2006), median interrater reliability (0.90; Roid, 2003), and concurrent validity (0.90 with SB-4, and 0.78 with Woodcock-Johnson III; Roid, 2003). The SB-5 also provides an Abbreviated Battery IQ (ABIQ), which is based on two routing subtests: One nonverbal (Object Series/Matrices) and one verbal (Vocabulary). Correlations between the ABIQ and full-scale IQ are strong (Twomey et al., 2018).

The Leiter International Performance Scale- 3rd Edition (Leiter-3). The Leiter International Performance Scale- 3rd Edition (Roid et al., 2013) is designed to assess nonverbal intellectual ability, memory, and attention in individuals aged 3 to 75 and above with two batteries: cognitive and attention/memory. The Cognitive Battery consists of five nonverbal subtests, four of which generate a Nonverbal IQ. This battery takes approximately 45 minutes to administer. No verbal instructions or responses are required, so this instrument can be used in testing non-speaking individuals with or without ASD. There was a small number of autistic individuals included in the standardization sample. The Leiter has been shown to have good reliability (0.91 - 0.93; Roid et al., 2009), acceptable to good test–retest reliability (0.70 - 0.88; Roid & Miller, 1997), and concurrent validity with other nonverbal measures (0.72; Hooper & Bell, 2006).

Adaptive Behavior

In addition to examining cognitive functioning, it is also important to assess adaptive behaviors, which typically consist of conceptual (e.g., concept of money, time), social (e.g., interpersonal skills), and practical skills (e.g., personal care, safety, transportation) that impact one's daily functioning (Schalock et al., 2010). Assessment of adaptive behavior helps determine individual strengths and challenges, inform treatment goals, and track response to intervention (Klinger et al., 2018; Ozonoff et al., 2005). In addition, to make a diagnosis of intellectual disability, the assessment of adaptive behavior is required (APA, 2020). The two most commonly used instruments assessing adaptive behaviors in autistic individuals are the Vineland Adaptive Behavior Scales (VABS; Sparrow et al., 2016) and Adaptive Behavior Assessment system (ABAS; Harrison & Oakland, 2015), which are both routinely used in clinical and research practices.

Vineland Adaptive Behavior Scales-Third Edition. The VABS-3 (Sparrow et al., 2016) has several forms that can be used to gather information from multiple sources (i.e., parent, teacher) regarding an individual's adaptive skills in the areas of social interaction and communication skills, personal living skills, community living skills, and motor skills across the lifespan (birth to 90 years). These include a survey interview form (administered by the examiner in an interview format), a questionnaire-based parent rating form, an expanded interview form, and a teacher rating form. Individuals with ASD were included in the standardization samples. The Vineland has good internal consistency reliability (0.77 - 0.93; de Bildt et al., 2005; Sparrow et al., 2005) with a test–retest reliability ranging from 0.53 to 0.80 (Sparrow et al., 2005) and good convergent validity (0.93; de Bildt et al., 2005). The concurrent validity between the Vineland and other adaptive scale is good (0.70 with ABAS; Perry & Factor, 1989; Sparrow et al., 2005).

Adaptive Behavior Assessment System- Third Edition. The ABAS-3 (Harrison & Oakland, 2015) consists of a collection of rating scales that examine adaptive skills in individuals 0 to 89 years of age. Different versions exist for different informants, including caregivers, teachers, or self-report. Autistic individuals were included in the standardization samples. The internal consistency reliability is good (0.85 - 0.99) with acceptable to good average test–retest reliability (0.70 - 0.89; Harrison & Oakland, 2015). The average correlation between the ABAS-3 parent form and Vineland-II parent form across skill areas is 0.66 (Harrison & Oakland, 2015). Similar to the considerations of choosing a measure in other domains, factors such as areas of information covered by the measure, age range, normative sample, informant (caregiver, teacher, self), details of the information needed (interview vs. checklist), and time of administration need to be considered when selecting a measure for cognitive and adaptive skills.

Other Assessment Considerations. The results of a comprehensive evaluation can help identify appropriate intervention goals and individualize treatment accordingly to provide patient- and family-centered care (Zwaigenbaum, Bishop, et al., 2021). A needs assessment is important as part of the diagnostic evaluation and treatment planning process to inform support and resources necessary for the individual to reach their full potential (Brian et al., 2019; Whitehouse et al., 2018). Other countries suggest other approaches. For example, the Australian national guidelines recommend conducting a comprehensive needs assessment as the first step in the process (Whitehouse et al., 2018), before the diagnostic evaluation. There are also other assessment considerations in the course of interventions to measure and assure progress toward goals. Instruments specifically designed for ongoing intervention planning and evaluation and to assess treatment progress and efficacy exist and have been validated, such as the Early Start Denver Model (ESDM) curriculum checklist (Rogers & Dawson, 2009), which assesses ASD-specific social and preverbal communication development and can be used to construct individualized treatment objectives (Contaldo et al., 2020; Rogers et al., 2014) and to track intervention progress (Contaldo et al., 2020).

Assessment of Co-Occurring Conditions in Autism

Co-occurring conditions (i.e., comorbidity) are common in autistic individuals. In an epidemiological, population-derived sample, more than 70% of autistic individuals had at least one co-occurring psychiatric condition, and more than 40% had two or more (Simonoff et al., 2008). Common comorbidities include anxiety disorders (20% vs. 7.3% in the general population), attention-deficit/hyperactivity disorder (ADHD, 28% vs. 7.2%), and depression (11% vs. 4.7%), according to a recent meta-analysis (Lai et al., 2019). According to large, community-based studies, the rate of intellectual disability in autism is estimated to be 35.2% (Maenner et al., 2021).

The potential for co-occurring conditions adds complexity to autism assessment. For example, avoidance and irritability are found in both autism and anxiety (Mayes et al., 2011). Diminished eye contact and low social initiative can be seen in autism and depression (J. A. Kim et al., 2000). Autism and ADHD are both associated with social challenges, language difficulties, sensory issues, and attention problems (Mayes et al., 2012). In addition, many psychiatric diagnostic instruments' standardization samples have failed to include autistic individuals or individuals with various levels of cognitive abilities (Deprey & Ozonoff, 2018). This makes the interpretation of assessment results a challenge. Furthermore, diagnosis of psychiatric conditions (e.g., anxiety, depression, obsessive–compulsive disorder) often relies on self-report of internal experience, which can be more challenging for autistic individuals, especially those with intellectual disabilities and/or limited language (Mazefsky et al., 2011).

Accurate assessment of behavioral and psychiatric concerns in individuals with autism allows for specific and individualized treatment planning and management of additional symptoms and in turn optimal treatment outcomes. When making differential diagnoses between autism and other conditions, and when considering comorbid diagnoses, it is important to evaluate the trajectory of symptom development (onset, duration, course). Co-occurring conditions should be considered when there are symptoms present which are not explained by autism, when there are changes in functioning from baseline, or when the individual is not responding to interventions (Deprey & Ozonoff, 2018). Additional assessment measures are often needed when concerns regarding co-occurring conditions are present, such as the Child Behavior Checklist (CBCL; Achenbach, 2001) and Behavioral Assessment System for Children-3 (Reynolds et al., 2015), which are two commonly used broadband measures of psychopathology, both of which have been used in samples of children with autism (e.g., Grondhuis & Aman, 2012; Guerrera et al., 2019; Stratis & Lecavalier, 2017). Notably, for the CBCL, reliability values have been smaller on average for autistic individuals than the original norming sample (Guerrera et al., 2019).

Future Directions in Autism Assessment

Telehealth and Beyond

In response to the COVID-19 pandemic, a number of adaptations were made to standard diagnostic assessment processes to meet family needs while also responding to physical distancing requirements present throughout the world. A shift to telehealth challenged providers in implementing comprehensive diagnostic assessment, necessitating modifications to standardized testing. This was particularly difficult for autism researchers and providers given that many of the gold-standard tools for evaluation rely on, among other things, the ability to communicate nonverbally, to observe eye contact and facial expressions, and to share affect. The use of personal protective equipment, such as face masks, raised many questions about the validity of these widely used tools, resulting in the development of modified or novel measures designed in response to these challenges. The impact of these modifications on the validity of the assessments remains to be fully determined, but efforts in that regard are underway.

One such tool, developed during the pandemic, is the *Brief Observation of Symptoms of Autism* (BOSA; Lord et al., 2020). This measure was developed to evaluate autism-related characteristics in a naturalistic social context without direct face-to-face contact between an individual and a clinician. Adapted from the Brief Observation of Social Communication Change (Grzadzinski et al., 2016) and ADOS-2 (Lord et al., 2012), the BOSA consists of a 12 to 14-minute interaction between a child and a caregiver, both unmasked. Providers observe behind an observation window and provide directions remotely to the caregiver using standardized materials and a series of standardized activities. The BOSA has four versions: Minimally Verbal (MV), Phrase Speech-Young Fluent (PSYF), Fluent Speech 1 (F1, 6 – 10 years), and Fluent Speech 2 (F2, 11-years and older). The BOSA-MV is for individuals who are nonverbal at any age, while the PSYF is for individuals with flexible phrase speech at any age or those who are verbally fluent under age 6. F1 and F2 are for younger and older verbally fluent individuals. Clinicians observe the caregiver-child interactions and then code the observation using ADOS-2 protocols, based on the individuals' age and language level (similar to the ADOS-2; e.g., for BOSA-PSYF, a Module 2 or 3 should be used). The scores then are transferred to a *DSM-5* checklist specific to that module and converted to binary BOSA scores, some of which contribute to a total algorithm score. Each module has a specified algorithm cutoff score for autism risk (to ensure around 90% sensitivity while maintaining adequate specificity).

The BOSA has been shown to have good discriminate validity between autism and non-spectrum groups, with acceptable-to-good sensitivity and specificity (sensitivity above 0.86 and specificity above 0.80, Dow et al., 2021). Good convergent validity was found between the BOSA and ADOS-2 in the Toddler Module and Module 3 (Dow et al., 2021). Interrater reliability was high, with intraclass correlation coefficients ranging from 0.90 to 0.95. Test–retest reliability was .95. Because it is a brief behavioral observation, to be able to make an autism diagnosis, gathering additional information is even more important (Lord et al., 2020). It is recommended by the authors that the BOSA be used as a *Level 2* screener to determine whether a full evaluation is warranted, or a diagnosis is clear, in combination with a caregiver interview, developmental history, and any additional behavioral observation (Lord et al., 2020).

Although the BOSA demonstrated its utility within the context of a pandemic, it was developed out of necessity and its accuracy remains to be fully examined and replicated. It has been suggested that the BOSA may be less accurate than the ADOS-2 (Dow et al., 2021), especially for adolescents and adults with more subtle symptoms who are referred for an autism evaluation for the first time. In addition, while the ADOS-2 is administered by trained clinicians who are skilled at eliciting social initiations and other behaviors from the child, this may not be the case when the BOSA is administered by parents. Caregivers, who are well-versed in the needs of their children, may be more inclined to scaffold for the child. Therefore, there may be fewer opportunities during the parent-administered BOSA to observe certain behaviors and challenges (e.g., joint attention, restricted and repetitive behaviors, conversation) both because of caregivers' tendency to accommodate to the child's needs and because of the limited time and activities included in the BOSA. Finally, because the BOSA requires a visit to a clinical setting and/or mailing a box of materials to the family, its accessibility and feasibility is limited.

Rather than being developed *in response* to the pandemic, the TELE-ASD-PEDS (TAP; Corona et al., 2020) is a tele-assessment instrument developed *before* the pandemic to address the research-to-practice gap in early diagnosis and resource burden in assessment services. It is designed to evaluate autism-related symptoms in young children younger than 36 months of age through observations of caregiver and child interactions. Caregivers use available toys and materials found in their environment to interact with the child, following instructions from the provider who observes via video-based web-conferencing. Providers coach caregivers to administer designated activities (e.g., interactive play, snack routine) and score based on child's behavior (e.g., socially directed speech and sounds, eye contact, repetitive play) using both dichotomous (yes/no) and Likert-type scales (3 = b*ehaviors characteristic of ASD clearly present*; 2 = *behaviors characteristic of ASD present at subclinical levels*; 1 = *behaviors characteristic of ASD not present*). The Likert-type ratings are summed, with a score of 12 or above indicative of autism risk. The TAP takes 15 to 20 min to administer. Current psychometric properties of the TAP are being investigated. Preliminary data (Wagner et al., 2021) suggested that most providers report feeling comfortable completing tele-assessments and making diagnoses of ASD following the TAP, and all providers reported the value of observing children and families in their home environments, relieving the burden of travel and transportation for families. Minor challenges, infrequently experienced, included technological issues, dropped calls, inconsistent audio, and distractions at home.

The Telehealth Evaluation of Development for Infants (TEDI; Talbott et al., 2019; Talbott, Dufek, et al., 2022) is a tele-assessment protocol adapted from laboratory-based measures of early autism characteristics (e.g., AOSI), communication, language, play, and other related developmental domains. It was developed prior to the pandemic to evaluate infants 6 to 12 months of age. The protocol involves a combination of caregiver questionnaires and a telehealth session, which consists of semi-structured parent-child play interactions, administered to the child by their primary caregiver under the instruction of a provider observing and interacting via telehealth platform. Providers score several measures (e.g., AOSI, the ESDM curriculum checklist, the Early Communication Indicator; Greenwood et al., 2010) based on the interaction. The initial study suggested acceptable interrater reliability (0.65; Talbott et al., 2019) and positive caregiver feedback (Talbott, Lang, et al., 2022).

The Parent-Administered Neurodevelopmental Assessment (Kelleher et al., 2020) is a tele-assessment instrument administered by caregivers that involves the collection of a wide array of information (e.g., attention, language, motor, stereotypical behaviors, physiological data) about infants and toddlers to assess early developmental concerns. Example tasks include passive viewing of a video, independent play, prompts to elicit autism-related symptoms, and parent–child interaction. Additional autism-specific presses by caregivers are under development. Preliminary data from a sample of infants with Down syndrome suggested high caregiver fidelity and satisfaction and infant engagement (Kelleher et al., 2020).

Other remote assessments have been developed in an effort to address the issues of long waiting times in obtaining autism diagnosis and the lack of workforce, and to improve equity in assessment access including the Systematic Observation of Red Flags of ASD (Dow et al., 2020), Naturalistic Observation Diagnostic Assessment (C. J. Smith et al., 2017), and the adaptations of the Childhood Autism Rating Scale (Schopler et al., 2010) rated by providers observing an interaction via telehealth platform. It is likely that some of these developments will be carried forward into standard practice and that there will be continued growth in this realm. Some states and clinics have already adopted hybrid assessment models allowing for combinations of telehealth and in-person visits (Jang et al., 2022).

In recent years, various technologies, such as mobile applications and artificial intelligence methods, have been explored in screening for and diagnosing autism. For example, ASDetect is a free mobile application for caregivers of children aged 11 to 30 months that can be used to assess the likelihood of autism based on caregiver-reported behavioral markers (Barbaro & Yaari, 2020). Other web or mobile applications have been developed for the early detection of autism through eye tracking (Chang et al., 2021) or caregiver-rated videos (Young et al., 2020). In addition, studies have explored using machine learning and artificial intelligence to aid in autism screening and diagnosis (Megerian et al., 2022; Shahamiri & Thabtah, 2020; Tariq et al., 2018). This is still a developing field, but some findings have been promising in terms of sensitivity for early detection of autism, although specificity and positive predictive value are often lower.

Equity in Autism Assessment

Recent events (pandemic, social injustices) have revealed long-standing disparities in our tools and systems. Although young children can be reliably diagnosed at age 2, the median age of autism diagnosis in the United States is still around age 4 (Maenner et al., 2021). Racial and ethnic differences in ASD prevalence have long been documented (Baio, 2014; Baio et al., 2018; Maenner et al., 2021). Although recent findings indicate that overall ASD prevalence is similar across racial and ethnic groups, in certain parts of the country (Maenner et al., 2021), Hispanic children remain less likely to be identified with autism than White or Black children, and Black autistic children are diagnosed with intellectual disability at much higher rates than White autistic children (Maenner et al., 2021). Studies have suggested various factors that may relate to diagnostic disparities in underrepresented groups or communities (e.g., less access to and use of services, lack of providers, long waiting list, stigma, and varied cultural interpretations of symptoms; Angell et al., 2018; K. A. Smith et al., 2020). However, the key among these barriers is the lack of culturally and linguistically sensitive and appropriate assessment instruments. Indeed, there is a great need to develop evidence-based, standardized assessment tools in languages other than English and with diverse standardization samples (and/or to adapt current instruments; Angell et al., 2018). While some measures have been translated into many other languages, many widely used screening and diagnostic tools were originally developed and validated using homogeneous standardization samples (i.e., predominately White and male; Navarro-Pardo et al., 2021), potentially building bias into the measures themselves. Culturally and linguistically responsive adaptations of ASD screening and diagnostic tools have been developed (Harris et al., 2014; Magaña & Smith, 2013; Vanegas & Abdelrahim, 2016) but additional efforts are needed, especially in terms of representing greater diversity in assessment standardization samples and in improving the validity of assessment instruments in underrepresented populations. Likewise, disparities with regard to sex have raised concerns (Fuss et al., 2018; Halladay et al., 2015). For example, autistic females are diagnosed later than males (Shattuck et al., 2009), are more likely to be missed or misdiagnosed (Bargiela et al., 2016; Fuss et al., 2018), and are more likely to be diagnosed with co-occurring psychiatric conditions (Jadav et al., 2022). As many autism assessment tools have been developed validated on primarily male samples, questions remain regarding whether the resulting norms apply to females.

Summary

Addressing the gap between initial signs of autism, referral for evaluation, and early diagnosis is critical for both short- and long-term outcomes of families and children and heavily relies on adequate assessment approaches. While recent events have revealed that families may benefit from telehealth options, further research is needed to examine the reliability and validity of remote assessment instruments. The acceptability, feasibility, and accessibility of such approaches also require further examination. Finally, the development of culturally and linguistically informed assessments is a critical avenue for future work with implications for both research and practice. Involving stakeholders in this work (e.g., caregivers, autistic adults, providers), especially those who work with or are from marginalized communities, is imperative.

Declaration of Conflicting Interests

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

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Yue Yu https://orcid.org/0000-0002-7417-0171

References

Achenbach, T. (2001). *Child Behavior Checklist (CBCL)*. Achenbach System of Empirically Based Assessment.

Angell, A. M., Empey, A., & Zuckerman, K. E. (2018). A review of diagnosis and service disparities among children with autism from racial and ethnic minority groups in the United States. *International Review of Research in Developmental Disabilities*, *55*, 145–180.

American Psychiatric Association. (1980). *Diagnostic and statistical manual of mental disorders* (3rd ed.).

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.).

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text rev.).

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). American Psychiatric Publishing.

American Psychiatric Association. (2022). *Diagnostic and statistical manual of mental disorders* (5th ed., text rev.). American Psychiatric Publishing.

Baio, J. (2014). Prevalence of autism spectrum disorder among children aged 8 years-autism and developmental disabilities monitoring network, 11 sites, United States, 2010. *MMWR Surveillance Summaries*, *63*(2), 1–21.

Baio, J., Wiggins, L., Christensen, D. L., Maenner, M. J., Daniels, J., Warren, Z., Kurzius-Spencer, M., Zahorodny, W., Robinson, C., Rosenberg White, T., Durkin, M. S., Imm, P., Nikolaou, L., Yeargin-Allsopp, M., Lee, L.-C., Harrington, R., Lopez, M., Fitzgerald, R. T., & . . .Dowling, N. F. (2018). Prevalence of autism spectrum disorder among children aged 8 years—Autism and developmental disabilities monitoring network, 11 sites, United States, 2014. *MMWR. Surveillance Summaries*, *27*, 1–23https://doi.org/10.15585/mmwr.ss6706a1

Barbaro, J., & Yaari, M. (2020). Study protocol for an evaluation of ASDetect–A mobile application for the early detection of autism. *BMC Pediatrics*, *20*(1), Article 21. https://doi.org/10.1186/S12887-019-1888-6/FIGURES/3

Bargiela, S., Steward, R., & Mandy, W. (2016). The experiences of late-diagnosed women with autism spectrum conditions: An investigation of the female autism phenotype. *Journal of Autism and Developmental Disorders*, *46*(10), 3281–3294.

Bayley, N., & Aylward, G. P. (2019). *Bayley Scales of Infant and Toddler Development fourth edition (Bayley-4)*. NCS Pearson.

Berument, S. K., Rutter, M., Lord, C., Pickles, A., & Bailey, A. (1999). Autism Screening Questionnaire: Diagnostic validity. *The British Journal of Psychiatry*, *175*(5), 444–451.

Bishop, S. L., Guthrie, W., Coffing, M., & Lord, C. (2011). Convergent validity of the Mullen Scales of Early Learning and the differential ability scales in children with autism spectrum disorders. *American Journal on Intellectual and Developmental Disabilities*, *116*(5), 331–343. https://doi.org/10.1352/1944-7558-116.5.331

Botha, M., Hanlon, J., & Williams, G. L. (2021). Does language matter? Identity-first versus person-first language use in autism research: A response to Vivanti. *Journal of Autism and Developmental Disorders*, 53, 870–878. https://doi.org/10.1007/S10803-020-04858-W

Brian, J. A., Zwaigenbaum, L., & Ip, A. (2019). Standards of diagnostic assessment for autism spectrum disorder. *Paediatrics & Child Health*, *24*(7), 444–451. https://doi.org/10.1093/PCH/PXZ117

Bryson, S. E., Zwaigenbaum, L., McDermott, C., Rombough, V., & Brian, J. (2008). The Autism Observation Scale for Infants: Scale development and reliability data. *Journal of Autism and Developmental Disorders*, *38*(4), 731–738.

Buijsman, R., Begeer, S., & Scheeren, A. M. (2022). "Autistic person" or "person with autism"? Person-first language preference in Dutch adults with autism and parents. *Autism*, *27*, 788–795. https://doi.org/10.1177/13623613221117914

Chang, Z., Di Martino, J. M., Aiello, R., Baker, J., Carpenter, K., Compton, S., Davis, N., Eichner, B., Espinosa, S., Flowers, J., Franz, L., Harris, A., Howard, J., Perochon, S., Perrin, E. M., Krishnappa Babu, P. R., Spanos, M., Sullivan, C., Walter, B. K., & . . .Sapiro, G. (2021). Computational methods to measure patterns of gaze in toddlers with autism spectrum disorder. *JAMA Pediatrics*, *175*(8), 827–836. https://doi.org/10.1001/JAMAPEDIATRICS.2021.0530

Charman, T., & Baird, G. (2002). Practitioner review: Diagnosis of autism spectrum disorder in 2-and 3-year-old children. *Journal of Child Psychology and Psychiatry, 43*(3), 289–305.

Chawarska, K., Klin, A., Paul, R., Macari, S., & Volkmar, F. (2009). A prospective study of toddlers with ASD: Short-term diagnostic and cognitive outcomes. *Journal of Child Psychology and Psychiatry*, *50*(10), 1235–1245.

Contaldo, A., Colombi, C., Pierotti, C., Masoni, P., & Muratori, F. (2020). Outcomes and moderators of Early Start Denver Model intervention in young children with autism spectrum disorder delivered in a mixed individual and group setting. *Autism*, *24*(3), 718–729. https://doi.org/10.1177/1362361319888344

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.

Dai, Y. G., Porto, K. S., Skapek, M., Barton, M. L., Dumont-Mathieu, T., Fein, D. A., & Robins, D. L. (2021). Comparison of the Modified Checklist for Autism in Toddlers, Revised with Follow-up (M-CHAT-R/F) positive predictive value by race. *Journal of Autism and Developmental Disorders*, *51*(3), 855–867. https://doi.org/10.1007/S10803-020-04428-0

Davidovitch, M., Levit-Binnun, N., Golan, D., & Manning-Courtney, P. (2015). Late diagnosis of autism spectrum disorder after initial negative assessment by a multidisciplinary team. *Journal of Developmental and Behavioral Pediatrics*, *36*(4), 227–234. https://doi.org/10.1097/DBP.0000000000000133

Dawkins, T., Meyer, A. T., & Van Bourgondien, M. E. (2016). The Relationship Between the Childhood Autism Rating Scale: Second edition and clinical diagnosis utilizing the DSM-IV-TR and the DSM-5. *Journal of Autism and Developmental Disorders*, *46*(10), 3361–3368. https://doi.org/10.1007/S10803-016-2860-Z/TABLES/5

de Bildt, A., Sytema, S., Kraijer, D., Sparrow, S., & Minderaa, R. (2005). Adaptive functioning and behaviour problems in relation to level of education in children and adolescents with intellectual disability. *Journal of Intellectual Disability Research*, *49*(Pt. 9), 672–681. https://doi.org/10.1111/j.1365-2788.2005.00711.x

Deprey, L., & Ozonoff, S. (2018). *Assessment of comorbid psychiatric conditions in autism spectrum disorder*.

DiLavore, P. C., Lord, C., & Rutter, M. (1995). The pre-linguistic autism diagnostic observation schedule. *Journal of Autism and Developmental Disorders*, *25*(4), 355–379. https://doi.org/10.1007/BF02179373

DiStefano, C., & Dombrowski, S. C. (2006). Investigating the theoretical structure of the Stanford-Binet. *Journal of Psychoeducational Assessment*, *24*(2), 123–136.

Dow, D., Day, T. N., Kutta, T. J., Nottke, C., & Wetherby, A. M. (2020). Screening for autism spectrum disorder in a naturalistic home setting using the systematic observation of red flags (SORF) at 18–24 months. *Autism Research*, *13*(1), 122–133.

Dow, D., Holbrook, A., Toolan, C., McDonald, N., Sterrett, K., Rosen, N., Kim, S. H., & 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*, 5383–5394. https://doi.org/10.1007/S10803-021-05395-W/TABLES/7

Elder, J. H., Kreider, C. M., Brasher, S. N., & Ansell, M. (2017). Clinical impact of early diagnosis of autism on the prognosis and parent-child relationships. In. *Psychology Research and Behavior Management*, *10*, 283–292. https://doi.org/10.2147/PRBM.S117499

Elliot, C. (2007). *Differential Abilities Scale—2nd edition (DAS-II) manual*. Harcourt Assessment.

Farmer, C., Golden, C., & Thurm, A. (2016). Concurrent validity of the differential ability scales, with the Mullen Scales of Early Learning in young children with and without neurodevelopmental disorders. *Child Neuropsychology*, *22*(5), 556–569.

Fuss, J., Briken, P., & Klein, V. (2018). Gender bias in clinicians' pathologization of atypical sexuality: A randomized controlled trial with mental health professionals. *Scientific Reports*, *8*(1), 1–9.

Gau, S. S. F., Lee, C. M., Lai, M. C., Chiu, Y. N., Huang, Y. F., Kao, J.-D., & Wu, Y. Y. (2011). Psychometric properties of the Chinese version of the Social Communication Questionnaire. *Research in Autism Spectrum Disorders*, *5*(2), 809–818. https://doi.org/10.1016/J.RASD.2010.09.010

Gernsbacher, M. A. (2017). Editorial perspective: The use of person-first language in scholarly writing may Accentuate stigma. *Journal of Child Psychology and Psychiatry*, *58*(7), 859–861. https://doi.org/10.1111/JCPP.12706

Gotham, K., Risi, S., Pickles, A., & Lord, C. (2007). The autism diagnostic observation schedule: Revised algorithms for improved diagnostic validity. *Journal of Autism and Developmental Disorders*, *37*(4), 613–627. https://doi.org/10.1007/S10803-006-0280-1/TABLES/6

Greenwood, C. R., Walker, D., & Buzhardt, J. (2010). The early communication indicator for infants and toddlers: Early Head Start growth norms from two states. *Journal of Early Intervention*, *32*(5), 310–334.

Grondhuis, S. N., & Aman, M. G. (2012). Assessment of anxiety in children and adolescents with autism spectrum disorders. In. *Research in Autism Spectrum Disorders*, *6*(4), 1345–1365. https://doi.org/10.1016/j.rasd.2012.04.006

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.

Guerrera, S., Menghini, D., Napoli, E., Di Vara, S., Valeri, G., & Vicari, S. (2019). Assessment of psychopathological comorbidities in children and adolescents with autism spectrum disorder using the Child Behavior Checklist. *Frontiers in Psychiatry*, 10, Article 535. https://doi.org/10.3389/fpsyt.2019.00535

Guo, C., Luo, M., Wang, X., Huang, S., Meng, Z., Shao, J., Zhang, X., Shao, Z., Wu, J., Robins, D. L., & Jing, J. (2019). Reliability and validity of the Chinese Version of Modified Checklist for Autism in Toddlers, Revised, with Follow-up (M-CHAT-R/F). *Journal of Autism and Developmental Disorders*, *49*(1), 185–196. https://doi.org/10.1007/S10803-018-3682-Y/TABLES/6

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

Guthrie, W., Wallis, K., Bennett, A., Brooks, E., Dudley, J., Gerdes, M., Pandey, J., Levy, S. E., Schultz, R. T., & Miller, J. S. (2019). Accuracy of autism screening in a large pediatric network. *Pediatrics*, *144*(4), Article e20183963.

Halladay, A. K., Bishop, S., Constantino, J. N., Daniels, A. M., Koenig, K., Palmer, K., Messinger, D., Pelphrey, K., Sanders, S. J., Singer, A. T., Taylor, J. L., & Szatmari, P. (2015). Sex and gender differences in autism spectrum disorder: Summarizing evidence gaps and identifying emerging areas of priority. *Molecular Autism*, *6*(1), Article 36. https://doi.org/10.1186/s13229-015-0019-y

Hardy, S., Orinstein, A., Barton, M., & Fein, D. (2013). Screening measures. In Volkmar, F. R. *(Ed.)*, *Encyclopedia of autism spectrum disorders* (pp. 2676–2682). Springer. https://doi.org/10.1007/978-1-4419-1698-3_607

Harris, B., Barton, E. E., & Albert, C. (2014). Evaluating autism diagnostic and screening tools for cultural and linguistic responsiveness. *Journal of Autism and Developmental Disorders*, *44*(6), 1275–1287.

Harrison, P. L., & Oakland, T. (2015). *Adaptive Behavior Assessment System (ABAS-3): Manual*. Pearson Clinical Assessment.

Hess, P. (2022). *DSM-5 revision tweaks autism entry for clarity*. Spectrum. https://doi.org/10.53053/JALY5388

Hooper, V. S., & Bell, S. M. (2006). Concurrent validity of the universal nonverbal intelligence test and the Leiter International Performance Scale–revised. *Psychology in the Schools*, *43*(2), 143–148.

Hyman, S. L., Levy, S. E., Myers, S. M., Kuo, D. Z., Apkon, C. 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., Brei, T., Davis, B. E., Lipkin, P. H., Norwood, K., Coleman, C., Mann, M., & . . .Paul, L. (2020). Identification, evaluation, and management of children with autism spectrum disorder. *Pediatrics*, *145*(1), Article e20193447. https://doi.org/10.1542/PEDS.2019-3447

Jadav, N., & Bal, V. H. (2022). Associations between co-occurring conditions and age of autism diagnosis:

Implications for mental health training and adult autism research. *Autism Research*, *15*(11), 2112–2125. https://doi.org/10.1002/AUR.2808

Jadav, N., Bal, V. H., & Vanessa Bal, C. H. (2022). Associations between co-occurring conditions and age of autism diagnosis: Implications for mental health training and adult autism research. *Autism Research*. https://doi.org/10.1002/AUR.2808

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

Johnson, S., & Marlow, N. (2006). Developmental screen or developmental testing? *Early Human Development*, *82*(3), 173–183. https://doi.org/10.1016/J.EARLHUMDEV.2006.01.008

Karaminis, T., & Stavrakaki, S. (2022). The psychometric properties of the Greek version of the Social Communication Questionnaire. *Autism Research*, *15*(9), 1768–1780. https://doi.org/10.1002/AUR.2790

Kelleher, B. L., Halligan, T., Witthuhn, N., Neo, W. S., Hamrick, L., & Abbeduto, L. (2020). Bringing the laboratory home: PANDABox telehealth-based assessment of neurodevelopmental risk in children. *Frontiers in Psychology*, *11*, Article 1634.

Kim, J. A., Szatmari, P., Bryson, S. E., Streiner, D. L., & Wilson, F. J. (2000). The prevalence of anxiety and mood problems among children with autism and Asperger syndrome. *Autism*, *4*(2), 117–132. https://doi.org/10.1177/1362361300004002002

Kim, S. H., & Lord, C. (2012). Combining information from multiple sources for the diagnosis of autism spectrum disorders for toddlers and young preschoolers from 12 to 47 months of age. *Journal of Child Psychology and Psychiatry*, *53*(2), 143–151.

Kim, S. H., Thurm, A., Shumway, S., & Lord, C. (2013). Multisite study of new Autism Diagnostic Interview-Revised (ADI-R) algorithms for toddlers and young preschoolers. *Journal of Autism and Developmental Disorders*, *43*(7), 1527–1538. https://doi.org/10.1007/s10803-012-1696-4

Klinger, L. G., Mussey, J. L., & O'Kelley, S. (2018). Assessment of intellectual functioning in autism spectrum disorder. In S. Goldstein & S. Ozonoff (Eds.), *Assessment of autism spectrum disorder* (pp. 215–262). Guilford Press.

Kuriakose, S. (2014). Concurrent validity of the WISC-IV and DAS-II in children with autism spectrum disorder. *Journal of Psychoeducational Assessment*, *32*(4), 283–294.

Lai, M. C., & Baron-Cohen, S. (2015). Identifying the lost generation of adults with autism spectrum conditions. *The Lancet Psychiatry*, *2*(11), 1013–1027. https://doi.org/10.1016/S2215-0366(15)00277-1

Lai, M. C., Kassee, C., Besney, R., Bonato, S., Hull, L., Mandy, W., Szatmari, P., & Ameis, S. H. (2019). Prevalence of co-occurring mental health diagnoses in the autism population: A systematic review and meta-analysis. *The Lancet Psychiatry*, *6*(10), 819–829. https://doi.org/10.1016/S2215-0366(19)30289-5

Landa, R. J., Gross, A. L., Stuart, E. A., & Faherty, A. (2013). Developmental trajectories in children with and without autism spectrum disorders: The first 3 years. *Child Development*, *84*(2), 429–442.

Lebersfeld, J. B., Swanson, M., Clesi, C. D., & O'Kelley, S. E. (2021). Systematic review and meta-analysis of the clinical utility of the ADOS-2 and the ADI-R in diagnosing autism spectrum disorders in children. *Journal of Autism and Developmental Disorders*, *51*(11), 4101–4114. https://doi.org/10.1007/s10803-020-04839-z

Lord, C., Dow, D., Holbrook, A., & Kim, S. (2020). *Brief Observation of Symptoms of Autism (BOSA)*. Western Psychological Services.

Lord, C., Rutter, M., DiLavore, P. C., & Risi, S. (1999). *Autism Diagnostic Observation Schedule (ADOS)*. Western Psychological Services.

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

Luyster, R., Gotham, K., Guthrie, W., Coffing, M., Petrak, R., Pierce, K., Bishop, S., Esler, A., Hus, V., Oti, R., Richler, J., Risi, S., & Lord, C. (2009). The Autism Diagnostic Observation Schedule–Toddler module: A new module of a standardized diagnostic measure for autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *39*(9), 1305–1320. https://doi.org/10.1007/S10803-009-0746-Z/TABLES/7

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. https://doi.org/10.15585/MMWR.SS7011A1

Magaña, S., & Smith, L. E. (2013). The use of the Autism Diagnostic Interview-Revised with a Latino population of adolescents and adults with autism. *Journal of Autism and Developmental Disorders*, *43*(5), 1098–1105.

Mandy, W. P. L., Charman, T., & Skuse, D. H. (2012). Testing the construct validity of proposed criteria for DSM-5 autism spectrum disorder. *Journal of the American Academy of Child & Adolescent Psychiatry*, *51*(1), 41–50. https://doi.org/10.1016/J.JAAC.2011.10.013

Mason, D., Stewart, G. R., Capp, S. J., & Happé, F. (2022). Older age autism research: A rapidly growing field, but still a long way to go. *Autism in Adulthood*, *4*(2), 164–172. https://doi.org/10.1089/AUT.2021.0041/SUPPL_FILE/SUPPL_DATA4.DOCX

Mayes, S. D., Calhoun, S. L., Mayes, R. D., & Molitoris, S. (2012). Autism and ADHD: Overlapping and discriminating symptoms. *Research in Autism Spectrum Disorders*, *6*(1), 277–285. https://doi.org/10.1016/J.RASD.2011.05.009

Mayes, S. D., Calhoun, S. L., Murray, M. J., Ahuja, M., & Smith, L. A. (2011). Anxiety, depression, and irritability in children with autism relative to other neuropsychiatric disorders and typical development. *Research in Autism Spectrum Disorders*, *5*(1), 474–485. https://doi.org/10.1016/J.RASD.2010.06.012

Mazefsky, C. A., Kao, J., & Oswald, D. P. (2011). Preliminary evidence suggesting caution in the use of psychiatric self-report measures with adolescents with high-functioning autism spectrum disorders. *Research in Autism Spectrum Disorders*, *5*(1), 164–174. https://doi.org/10.1016/j.rasd.2010.03.006

Megerian, J. T., Dey, S., Melmed, R. D., Coury, D. L., Lerner, M., Nicholls, C. J., Sohl, K., Rouhbakhsh, R., Narasimhan, A., Romain, J., Golla, S., Shareef, S., Ostrovsky, A., Shannon, J., Kraft, C., Liu-Mayo, S., Abbas, H., Gal-Szabo, D. E., Wall, D. P., & Taraman, S. (2022). Evaluation of an artificial intelligence-based medical device for diagnosis of autism spectrum disorder. *NPJ Digital Medicine*, *5*(1), 57. https://doi.org/10.1038/S41746-022-00598-6

Michelotti, J., Psych, C., Centre, D. W., Charman, T., London, C., Slonims, V., & Baird, G. (2002). Follow-up of children with language delay and features of autism from preschool years to middle childhood. *Developmental Medicine and Child Neurology*, *44*(12), 812–819. https://doi.org/10.1017/S0012162201002985

Monaghan, T. F., Rahman, S. N., Agudelo, C. W., Wein, A. J., Lazar, J. M., Everaert, K., & Dmochowski, R. R. (2021). Foundational statistical principles in medical research: Sensitivity, specificity, positive predictive value, and negative predictive value. Medicina, *57*, 503. https://doi.org/10.3390/medicina57050503

Moon, S. J., Hwang, J. S., Shin, A. L., Kim, J. Y., Bae, S. M., Sheehy-Knight, J., & Kim, J. W. (2019). Accuracy of the Childhood Autism Rating Scale: A systematic review and meta-analysis. *Developmental Medicine & Child Neurology*, *61*(9), 1030–1038. https://doi.org/10.1111/DMCN.14246

Moulton, E., Bradbury, K., Barton, M., & Fein, D. (2019). Factor analysis of the childhood autism rating scale in a sample of two year olds with an autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *49*(7), 2733–2746.

Mullen, E. M. (1995). *Mullen Scales of Early Learning*. AGS.

Navarro-Pardo, E., López-Ramón, M. F., Alonso-Esteban, Y., & Alcantud-Marín, F. (2021). Diagnostic tools for autism spectrum disorders by gender: Analysis of current status and future lines. *Children*, *8*(4), 262. https://doi.org/10.3390/CHILDREN8040262

Norris, M., & Lecavalier, L. (2010). Screening accuracy of Level 2 autism spectrum disorder rating scales. A review of selected instruments. *Autism: The International Journal of Research and Practice*, *14*(4), 263–284. https://doi.org/10.1177/1362361309348071

Ozonoff, S., Goodlin-Jones, B. L., & Solomon, M. (2005). Evidence-based assessment of autism spectrum disorders in children and adolescents. *Journal of Clinical Child and Adolescent Psychology*, *34*(3), 523–540. https://doi.org/10.1207/s15374424jccp3403_8

Ozonoff, S., Iosif, A. M., Baguio, F., Cook, I. C., Hill, M. M., Hutman, T., & Steinfeld, M. B. (2010). A prospective study of the emergence of early behavioral signs of autism. *Journal of the American Academy of Child & Adolescent Psychiatry*, *49*(3), 256–266. https://doi.org/10.3816/CLM.2009.n.003.Novel

Ozonoff, S., Young, G. S., Brian, J., Charman, T., Shephard, E., Solish, A., & Zwaigenbaum, L. (2018). Diagnosis of autism spectrum disorder after age 5 in children evaluated longitudinally since infancy. *Journal of the American Academy of Child & Adolescent Psychiatry*, *57*(11), 849–857. https://doi.org/10.1016/j.jaac.2018.06.022

Ozonoff, S., Young, G. S., Landa, R. J., Brian, J., Bryson, S., Charman, T., Chawarska, K., Macari, S. L., Messinger, D., Stone, W. L., Zwaigenbaum, L., & Iosif, A. M. (2015). Diagnostic stability in young children at risk for autism spectrum disorder: A baby siblings research consortium study. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *56*(9), 988–998. https://doi.org/10.1111/jcpp.12421

Parikh, C., Iosif, A.-M., & Ozonoff, S. (2021). Brief report: Use of the Infant–Toddler Checklist in infant siblings of children with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *51*(3), 1007–1012.

Penner, M., Anagnostou, E., & Ungar, W. J. (2018). Practice patterns and determinants of wait time for autism spectrum disorder diagnosis in Canada. *Molecular Autism*, *9*(1), 1–13. https://doi.org/10.1186/S13229-018-0201-0/TABLES/8

Perry, A., & Factor, D. C. (1989). Psychometric validity and clinical usefulness of the Vineland Adaptive Behavior Scales and the AAMD Adaptive Behavior Scale for an autistic sample. *Journal of Autism and Developmental Disorders*, *19*(1), 41–55.

Petrocchi, S., Levante, A., & Lecciso, F. (2020). Systematic review of Level 1 and Level 2 Screening tools for autism spectrum disorders in toddlers. *Brain Sciences*, *10*(3), 180. https://doi.org/10.3390/BRAINSCI10030180

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

Piven, J., & Rabins, P. (2011). Autism spectrum disorders in older adults: Toward defining a research agenda. *Journal of the American Geriatrics Society*, *59*(11), 2151–2155. https://doi.org/10.1111/J.1532-5415.2011.03632.X

Powell, K. K., Heymann, P., Tsatsanis, K. D., & Chawarska, K. (2018). Assessment and diagnosis of infants and toddlers with autism spectrum disorder. In S. Goldstein & S. Ozonoff (Eds.), *Assessment of autism spectrum disorders* (pp. 96–129). Guilford.

Reynolds, C. R., Kamphaus, R. W., & Vannest, K. J. (2015). *BASC3: Behavior Assessment System for Children*. PscyhCorp.

Risi, S., Lord, C., Gotham, K., Corsello, C., Chrysler, C., Szatmari, P., Cook, E. H., Leventhal, B. L., & Pickles, A. (2006). Combining information from multiple sources in the diagnosis of autism spectrum disorders. *Journal of the American Academy of Child & Adolescent Psychiatry*, *45*(9), 1094–1103. https://doi.org/10.1097/01.chi.0000227880.42780.0e

Robins, D. L., Casagrande, K., Barton, M., Chen, C. M. A., Dumont-Mathieu, T., & Fein, D. (2014). Validation of the Modified Checklist for Autism in Toddlers, Revised with Follow-up (M-CHAT-R/F). *Pediatrics*, *133*(1), 37–45. https://doi.org/10.1542/peds.2013-1813

Robison, J. E. (2019). Talking about autism—Thoughts for researchers. *Autism Research*, *12*(7), 1004–1006. https://doi.org/10.1002/AUR.2119

Rogers, S. J., Estes, A., Lord, C., Vismara, L., Winter, J., Fitzpatrick, A., Guo, M., & Dawson, G. (2012). Effects of a brief early Start Denver Model (ESDM)-based parent intervention on toddlers at risk for autism spectrum disorders: A randomized controlled trial. *Journal of the American Academy of Child & Adolescent Psychiatry*, *51*(10), 1052–1065. https://doi.org/10.1016/j.jaac.2012.08.003

Rogers, S. J., Vismara, L., Wagner, A. L., McCormick, C., Young, G., & Ozonoff, S. (2014). Autism treatment in the first year of life: A pilot study of infant start, a parent-implemented intervention for symptomatic infants. *Journal of Autism and Developmental Disorders*, *44*(12), 2981–2995. https://doi.org/10.1007/s10803-014-2202-y

Rogers, S. J., & Dawson, G. (2009). *Early Start Denver Model curriculum checklist for young children with autism*. Guilford Press.

Roid, G. H. (2003). *Stanford-Binet Intelligence Scales (SB-5)*. Riverside publishing.

Roid, G. H., & Miller, L. J. (1997). *Leiter International Performance Scale-Revised (Leiter-R)*. Stoelting.

Roid, G. H., Miller, L. J., & Koch, C. (2013). *Leiter International Performance Scale*. Stoelting.

Roid, G. H., Pomplun, M., & Martin, J. J. (2009). Nonverbal intellectual and cognitive assessment with the Leiter International Performance Scale-Revised (Leiter-R). In J. A. Naglieri & S. Goldstein (Eds.), *Practitioner's guide to assessing intelligence and achievement* (pp. 265–290). John Wiley.

Rotholz, D. A., Kinsman, A. M., Lacy, K. K., & Charles, J. (2017). Improving early identification and intervention for children at risk for autism spectrum disorder. *Pediatrics*, *139*(2), Article e20161061. https://doi.org/10.1542/PEDS.2016-1061

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

Schalock, R. L., Borthwick-Duffy, S. A., Bradley, V. J., Buntinx, W. H. E., Coulter, D. L., Craig, E. M., Gomez, S. C., Lachapelle, Y., Luckasson, R., & Reeve, A. (2010). *Intellectual disability: Definition, classification, and systems of supports*. ERIC.

Schopler, E., Van Bourgondien, M. E., Wellman, G. J., & Love, S. R. (2010). *CARS-2: Childhood Autism Rating Scale–Second edition*. Western Psychological Services.

Shahamiri, S. R., & Thabtah, F. (2020). Autism AI: A new autism screening system based on artificial intelligence. *Cognitive Computation*, *12*(4), 766–777. https://doi.org/10.1007/S12559-020-09743-3/TABLES/4

Shattuck, P. T., Durkin, M., Maenner, M., Newschaffer, C., Mandell, D. S., Wiggins, L., Lee, L.-C., Rice, C., Giarelli, E., & Kirby, R. (2009). Timing of identification among children with an autism spectrum disorder: Findings from a population-based surveillance study. *Journal of the American Academy of Child & Adolescent Psychiatry*, *48*(5), 474–483.

Simonoff, E., Pickles, A., Charman, T., Chandler, S., Loucas, T., & Baird, G. (2008). Psychiatric disorders in children with autism spectrum disorders: Prevalence, comorbidity, and associated factors in a population-derived sample. *Journal of the American Academy of Child & Adolescent Psychiatry*, *47*(8), 921–929. https://doi.org/10.1097/CHI.0b013e318179964f

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. https://doi.org/10.1037/PAS0000317

Smith, K. A., Gehricke, J.-G., Iadarola, S., Wolfe, A., & Kuhlthau, K. A. (2020). Disparities in service use among children with autism: A systematic review. *Pediatrics*, *145*(Suppl. 1), S35–S46.

Sparrow, S. S., Cicchetti, D. V., & Balla, D. A. (2005). *Vineland Adaptive Behavior Scales Vineland-II: Survey forms manual*. Pearson.

Sparrow, S. S., Cicchetti, D. V., & Saulnier, C. A. (2016). *Vineland-3: Vineland Adaptive Behavior Scales*. PsychCorp.

Steiner, A. M., Koegel, L. K., Koegel, R. L., & Ence, W. A. (2012). Issues and theoretical constructs regarding parent education for autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *42*(6), 1218–1227. https://doi.org/10.1007/s10803-011-1194-0

Stone, W. L., Coonrod, E. E., Turner, L. M., & Pozdol, S. L. (2004). Psychometric properties of the STAT for early autism screening. *Journal of Autism and Developmental Disorders*, *34*(6), 691–701. https://doi.org/10.1007/s10803-004-5289-8

Stratis, E. A., & Lecavalier, L. (2017). Predictors of parent–teacher agreement in youth with autism spectrum disorder and their typically developing siblings. *Journal of Autism and Developmental Disorders*, *47*(8), 2575–2585.

Sutera, S., Pandey, J., Esser, E. L., Rosenthal, M. A., Wilson, L. B., Barton, M., Green, J., Hodgson, S., Robins, D. L., & Dumont-Mathieu, T. (2007). Predictors of optimal outcome in toddlers diagnosed with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *37*(1), 98–107.

Swineford, L. B., Guthrie, W., & Thurm, A. (2015). Convergent and divergent validity of the Mullen Scales of Early Learning in young children with and without autism spectrum disorder. *Psychological Assessment*, *27*(4), 1364–1378.

Talbott, M. R., Dufek, S., Young, G., & Rogers, S. J. (2022). Leveraging telehealth to evaluate infants with prodromal autism spectrum disorder characteristics using the Telehealth Evaluation of Development for Infants. *Autism*, *26*(5), 1242–1254.

Talbott, M. R., Dufek, S., Zwaigenbaum, L., Bryson, S., Brian, J., Smith, I. M., & Rogers, S. J. (2019). Brief report: Preliminary feasibility of the TEDI: A novel parent-administered telehealth assessment for autism spectrum disorder symptoms in the first year of life. *Journal of Autism and Developmental Disorders*, *50*(9), 3432–3439. https://doi.org/10.1007/s10803-019-04314-4

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.

Tariq, Q., Daniels, J., Schwartz, J. N., Washington, P., Kalantarian, H., & Wall, D. P. (2018). Mobile detection of autism through machine learning on home video: A development and prospective validation study. *PLOS Medicine*, *15*(11), Article e1002705. https://doi.org/10.1371/JOURNAL.PMED.1002705

Twomey, C., O'Connell, H., Lillis, M., Tarpey, S. L., & O'Reilly, G. (2018). Utility of an abbreviated version of the Stanford-Binet Intelligence Scales in estimating "full scale" IQ for young children with autism spectrum disorder. *Autism Research*, *11*(3), 503–508.

Vanegas, S. B., & Abdelrahim, R. (2016). Characterizing the systems of support for families of children with disabilities: A review of the literature. *Journal of Family Social Work*, *19*(4), 286–327. https://doi.org/10.1080/10522158.2016.1218399

Van Niekerk, M. E. H., Groen, W., Vissers, C. T. W. M., Van Driel-De Jong, D., Kan, C. C., & Oude Voshaar, R. C. (2011). Diagnosing autism spectrum disorders in elderly people. *International Psychogeriatrics*, *23*(5), 700–710. https://doi.org/10.1017/S1041610210002152

Vivanti, G. (2020). Ask the editor: What is the most appropriate way to talk about individuals with a diagnosis of autism? *Journal of Autism and Developmental Disorders*, *50*, 691–693. https://doi.org/10.1007/s10803-019-04280-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. (2021). 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

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

Wechsler, D. (2012). *Wechsler Preschool and Primary Scale of Intelligence—Fourth edition*. The Psychological Corporation.

Wechsler, D. (2014). *Wechsler Intelligence Scale for Children–Fifth edition (WISC-V)*. Pearson.

Werner, E., & Dawson, G. (2005). Validation of the phenomenon of autistic regression using home videotapes. *Archives of General Psychiatry*, *62*(8), 889–895. https://doi.org/10.1001/archpsyc.62.8.889

Wetherby, A. M., Allen, L., Cleary, J., Kublin, K., & Goldstein, H. (2002). Validity and reliability of the communication and symbolic behavior scales developmental profile with very young children. *Journal of Speech, Language, and Hearing Research, 45*(6), 1202–1218.

Wetherby, A. M., Brosnan-Maddox, S., Peace, V., & Newton, L. (2008). Validation of the Infant—Toddler Checklist as a broadband screener for autism spectrum disorders from 9 to 24 months of age. *Autism*, *12*(5), 487–511.

Whitehouse, A., Evans, K., Eapen, V., & Wray, J. (2018). *A national guideline for the assessment and diagnosis of autism spectrum disorders in Australia responses to public consultation submissions*. https://www.autismcrc.com.au/

World Health Organization. (2013, September 16–18). Meeting report: autism spectrum disorders and other developmental disorders: from raising awareness to building capacity, World Health Organization, Geneva, Switzerland.

Woolfenden, S., Sarkozy, V., Ridley, G., & Williams, K. (2012). A systematic review of the diagnostic stability of autism spectrum disorder. *Research in Autism Spectrum Disorders*, *6*(1), 345–354. https://doi.org/10.1016/j.rasd.2011.06.008

Young, G. S., Constantino, J. N., Dvorak, S., Belding, A., Gangi, D., Hill, A., Hill, M., Miller, M., Parikh, C., Schwichtenberg, A. J., Solis, E., & Ozonoff, S. (2020). A video-based measure to identify autism risk in infancy. *Journal of Child Psychology and Psychiatry*, *61*(1), 88–94. https://doi.org/10.1111/JCPP.13105

Zander, E., Willfors, C., Berggren, S., Coco, C., Holm, A., Jifält, I., Kosieradzki, R., Linder, J., Nordin, V., Olafsdottir, K., & Bölte, S. (2017). The interrater reliability of the Autism Diagnostic Interview-Revised (ADI-R) in clinical settings. *Psychopathology*, *50*(3), 219–227. https://doi.org/10.1159/000474949

Zwaigenbaum, L., Bishop, S., Stone, W. L., Ibanez, L., Halladay, A., Goldman, S., Kelly, A., Klaiman, C., Lai, M. C., Miller, M., Saulnier, C., Siper, P., Sohl, K., Warren, Z., & Wetherby, A. (2021). Rethinking autism spectrum disorder assessment for children during COVID-19 and beyond. *Autism Research*, *14*(11), 2251–2259. https://doi.org/10.1002/aur.2615

Zwaigenbaum, L., Bryson, S. E., Brian, J., Smith, I. M., Sacrey, L., Armstrong, V., Roberts, W., Szatmari, P., Garon, N., Vaillancourt, T., & Roncadin, C. (2021). Assessment of autism symptoms from 6 to 18 months of age using the Autism Observation Scale for infants in a prospective high-risk cohort. *Child Development*, *92*(3), 1187–1198. https://doi.org/10.1111/CDEV.13485

Zwaigenbaum, L., & Penner, M. (2018). Autism spectrum disorder: Advances in diagnosis and evaluation. *BMJ*, *361*, Article k1674. https://doi.org/10.1136/bmj.k1674