ORIGINAL ARTICLE

CLEAR Autism Diagnostic Evaluation (CADE): Evaluation of Reliability and Validity

M. Franci Crepeau-Hobson[1] [ORCID] · **Nancy Leech[1]** · **Courtney Russell[1]**

Abstract

Recent surveillance efforts indicate that 1 in 54 American children meet the criteria for Autism Spectrum Disorder (ASD), making it the fastest growing neurodevelopmental disorder in the U.S. Despite evidence that ASD can be reliably diagnosed as early as 24 months, the median age at ASD diagnosis in 2016 in the U.S. was 51 months. The CLEAR Autism Diagnostic Evaluation (CADE; Willard & Kroncke, 2019), was developed in response to the need to improve, shorten, and standardize the clinical ASD evaluation process. The CADE is a 33-item rating scale designed to be completed by caregivers and clinicians. The current study was conducted to examine the reliability and validity of the CADE using a sample of 191 individuals who received a private evaluation for ASD. Using the client's evaluation records, clinicians completed the CADE items. The coefficient alpha was .94, which indicates that the items form a scale that has high internal consistency. The CADE total scores were highly correlated with ADOS scores, with r values ranging from .52—.86, and discriminated between those participants with a diagnosis of ASD and those without ($p < .001$). Receiver operator characteristic (ROC) curve analyses indicated excellent diagnostic accuracy of the CADE total score (ROC area under the curve = .998). Results suggest that the CADE can be used as an efficient and accurate means of evaluating ASD. Limitations and implications for use of the CADE are discussed.

Keywords Autism spectrum disorder · Autism assessment · Diagnosis · Reliability · Validation

The most recent data published by the Autism and Developmental Disabilities Monitoring (ADDM) Network of the Centers for Disease Control and Prevention (CDC) indicates the prevalence of autism spectrum disorder (ASD) has increased to 18.5

✉ M. Franci Crepeau-Hobson
   franci.crepeau-hobson@ucdenver.edu

[1] School of Education and Human Development, University of Colorado Denver, Denver, CO, USA

per 1,000 (one in 54) children (Maenner et al., 2020). This is considerably higher than the previous ADDM Network ASD prevalence estimate of 16.8 per 1,000 (one in 59) children in 2014 (Baio et al., 2018) and approximately 2.5 times higher than the first ADDM Network ASD prevalence estimates of 6.7 (one in 150) from 2000 and 2002 (Maenner et al., 2020). The 2016 data indicate that ASD was 4.3 times as prevalent among boys as among girls.

ASD is characterized by two major clusters of symptoms: (1) deficits in social communication and interactions with others, and (2) repetitive or restricted behaviors and interests (American Psychology Association; APA, 2013). A systematic review of both retrospective studies and prospective studies of high-risk infants by Zwaigenbaum et al. (2013) provides strong evidence that ASD symptoms emerge in the first one to two years of life, and there is evidence that ASD can be reliably diagnosed as early as 24 months (Johnson & Myers, 2007; Lord et al., 2006). However, the ADDM data indicate that in 2016, the median age at ASD diagnosis was 51 months with a range of 38 months to 57 months. Children with co-occurring intellectual disability were diagnosed much earlier on average than those with IQ's of > 70 (44 months vs. 57 months). Research has highlighted delayed diagnosis among Hispanic children (Daniels & Mandell, 2013; Maenner et al., 2020), families with low SES, and those living in under-resourced or rural areas (Daniels & Mandell, 2013).

The relatively late diagnosis of ASD is of concern given that early identification and intensive early intervention are associated with more favorable outcomes for children with ASD (Fuller & Kaiser, 2019; Rogers & Vismara, 2008). A recent meta-analysis conducted by Fuller and Kaiser (2019) indicated that social communication can be markedly improved by early intervention with the largest effect sizes for studies that had an average age of participants of 3.81 years. Similarly, improved treatment progress and reduction of ASD-related symptoms have been specifically linked to age of entry into intervention with younger children (ages 2.5–5.15) making significantly more progress in their behavioral intervention programs compared with children ages 5.15 to 7.14 years (Itzchak & Zachor, 2011). Additionally, earlier diagnosis of ASD can impact school placement, with earlier identification and placement in services associated with placement in more inclusive school settings (Harris & Handleman, 2000). Clearly, there is empirical support of the need for increased access to intervention at earlier ages. However, access to ASD specific intervention and funding services is largely dependent on a diagnosis of the disorder (Bent et al., 2017; Hyman et al., 2020).

There are several significant barriers that commonly delay or prevent access to an ASD diagnosis. One potential reason for this delay is the time it takes to get an evaluation that leads to an ASD diagnosis. A study by Bisgaier and colleagues (Bisgaier et al., 2011) examined wait times to get an appointment for an evaluation for ASD. They found an average wait time of 85 days for U.S. families with Medicaid/CHIP–enrolled children and 77 days for families with health insurance. Additionally, more than half (57%) of the 14 clinics in the study required the families complete a lengthy clinical screening form (average of 158 questions) prior to scheduling appointments (Bisgaier et al., 2011). Studies examining time to diagnosis also found long delays. For example, researchers in the United Kingdom

discovered an average delay of 3.5 years for an ASD diagnosis after the initial contact with a health care professional by parents (Crane et al., 2016). Similar delays have been noted in Canada. Siklos and Kerns (2007) found that it took almost three years and visits to an average of 4.5 professionals to receive a diagnosis following their first visit to a professional regarding their child's development. Even when a child has been evaluated, there still may be a delay in getting the actual ASD diagnosis, In a population-based study conducted in the U.S., the average delay between the average age at first documented evaluation and the average age at first ASD diagnosis was 13 months (Wiggins et al., 2006).

McMorris et al. (2013) posited that challenges in diagnosing ASD among health professionals may contribute to delays in diagnosis as parents are required to visit multiple clinicians before receiving a diagnosis for their child. These data are suggestive of a critical shortage of clinicians who are trained in ASD diagnosis. This may be related in part to the extensive training requirements of widely used ASD evaluation tools.

Beyond a delay in diagnosis, the process is associated with significant parent dissatisfaction and increased stress and feelings of parent burnout (Crane et al., 2016; Siklos & Kerns, 2007). In addition to long wait times, the actual evaluation process can be expensive and quite lengthy. From initial consultation and testing to report writing, the clinical ASD evaluation and diagnostic process can last months, with one study finding psychologists spending an average of 152 days completing evaluations (Ahlers et al., 2019). This is not surprising, given the current recommendation for ASD evaluations to include a lengthy structured parent interview and professional observations of the child's behavior via a standardized observation assessment (Risi et al., 2006).

One of the most widely used and highly recommended observation instruments for ASD assessment is the Autism Diagnostic Observation Schedule (ADOS; Lord et al., 2000). The ADOS is a 40- to 60-min, play-based, standardized assessment for individuals ages 12 months and up. The most recent version of the ADOS, the ADOS-2 (Lord et al., 2012) includes four module administration options depending on the language abilities of the client. Diagnostic decisions are based on algorithm items. The ADOS-2 is a complex diagnostic instrument with respect to administration and coding of observed behavior, and it requires intensive training to accurately administer and score (Hyman et al., 2020) and routine use of the test to maintain efficiency in administration and coding. The high cost, time, and demanding training required for standardized direct observations such as the ADOS likely contribute to the insufficient supply of clinicians available to diagnose ASD and the resulting wait times and delays in diagnosis and treatment.

Clearly, the current diagnostic assessment process is failing to accommodate the increase in prevalence rate of ASD, overburdens families, and fails to provide children with access opportunities to improve their outcomes (Willard et al., 2019). Moreover, the burden of waiting impacts familial quality of life and also decreases satisfaction with the healthcare process. When told to "wait and see," families report uncertainty, frustration and decreased satisfaction with the current process (Lappé et al., 2018). Addressing these challenges requires a more efficient diagnostic process. In response to this need, a dynamic process, the CLEAR Autism Diagnostic

Evaluation (CADE; Willard & Kroncke, 2019), was developed to aid clinicians in the diagnosis of ASD. The CADE is designed to improve, shorten, and standardize the clinical ASD evaluation process while maintaining accuracy and providing support for clinical judgement. The current study was designed to evaluate the reliability and validity of the CADE in a sample of individuals who were evaluated for ASD.

Development and Description of the CADE

The CLEAR Autism Diagnostic Evaluation (CADE) scale was developed by Willard and Kroncke (2019) prior to the current study. The development of the CADE scale followed the recommended phases as described a recent peer-reviewed article (Boateng et al., 2018). These phases include 1) items are generated and the validity of their content is assessed; 2) the scale is constructed (i.e., pre-test questions, administer survey, reduce number of items, understand number of factors captured); and 3); scale evaluation.

The CADE Development process began with content identification, as well as the refinement of a framework developed over 3 years in clinical practice that identified 11 child development domains (Kroncke et al., 2016). This framework was used to generate the initial survey items. Concurrent with the initial development of the CADE, 44 stakeholders were interviewed. These stakeholders were occupational therapists, physical therapists, school counselors, school psychologists, clinical psychologists, pediatricians, developmental pediatricians, psychiatrists, researchers, parents, and teachers. The purpose of the interviews was to capture the lived experience of those who had encountered the lack of resources associated with the identification and diagnosis of ASD. The authors sought their input regarding the kind of tool would be most helpful to them in the process of identifying ASD, as well as in collaborating on treatment.

Content validity was assessed in three rounds via a panel of 13 subject matter experts. The panel included school and clinical psychologists, as well as developmental pediatricians; each of these clinicians had expertise in ASD. These rounds included a review of the 11 domains to determine if CADE items adequately addressed them; interviews with these subject matter experts to assess whether the items on the scale appropriately captured the content represented by each item; and expert ratings of the items for clarity, specificity, and sensitivity. Item ratings by the experts could range from 0 to 3, with 0 indicating no clarity, specificity, or sensitivity and 3 indicating the item showed clarity, specificity, and sensitivity. Items rated less than 2.8 by the experts were either deleted or revised. Revised items were rated by additional subject matter experts and were included only if rated 2.8 or greater by all experts.

The initial draft of the CADE rating scale included 104 items that were derived from the original 11 domains of assessment. The items were first reviewed by the CADE authors for diagnostic relevance to ASD. Any items that were deemed to be

poorly worded by the developers of CADE were reviewed and adjusted before completion of the original draft of 104 items.

The 13 original subject matter experts provided additional input to refine the CADE by refining item wording and evaluating item coverage across the ASD domains, as well as alignment with symptoms assessed by the ADOS, and with DSM-5 diagnostic criteria. Following several rounds of review, evaluation of clarity, specificity and sensitivity, and cognitive interviews, the final version of the scale included 80 items. Of these, 33 contribute to an ASD Total Score. It is this score that is aimed at facilitating a quicker and more accurate diagnosis of ASD. The other 47 items assess various areas of an individual's life, such as adaptive living skills, behavior, cognitive functioning, motor skills, learning, and memory and are utilized for treatment planning only. Thus, skills are rated across developmental areas such that autism can be diagnosed, and any challenging or troubling symptoms that impact functioning in life can be identified. This process allows for recommendations to be clearer and more targeted.

The CADE is designed to be completed by the parent or primary caregiver prior to the initial consultation with the clinician and is subsequently completed by the clinician using all available data. The parent is asked to identify two additional raters such as a teacher, physician, therapist, or other service provider who can be asked to rate the client using the CADE as well. Although these additional ratings are optional, having ratings from at least one other individual is preferred for educational and treatment planning. Using the data from the rater(s), as well as their diagnostic impressions during the evaluation, the clinician also completes the CADE.

Method

This study investigated the following research questions:

1) Does the CADE demonstrate adequate internal consistency?
2) Do CADE Total scores differ based on gender or race/ethnicity?
3) Are CADE Total scores significantly associated with ADOS scores?
4) What is the ASD diagnostic accuracy of CADE total scores?
5) Do CADE Total scores differ between individuals with a diagnosis of ASD and those who do not have a diagnosis?
6) Do CADE Total scores differ between individuals with differing levels of severity of ASD as rated by the clinician?
7) Do CADE Total scores differ between individuals with a Full Scale IQ score in the Intellectual Disability range (< 70) and those with Full Scale IQs above 70?
8) Are CADE Total scores associated with Wechsler Scale Full Scale IQ Scores?

Data Sources and Collection Procedures

Recruited from an ASD clinician listserv in Colorado, a total of five raters were selected for CADE scale validation. All but one were doctoral-level, licensed clinical

psychologists who are engaged in ASD diagnostic work on a daily basis. The fifth rater was a doctoral-level school psychologist working at a local private school for children with autism. Raters' years of experience in ASD assessment ranged from 5 to 20 years. Four raters were female and one was male. All of these clinicians were known locally in the field as specializing in ASD assessment. Each rater had recently completed an ASD assessment course and each was provided CADE-specific training including detailed descriptions of each item and each level of scoring. All raters received a CADE manual with detailed descriptions and scoring guidelines for each item. The CADE developers were available to provide support as needed for the raters. Raters received a stipend of $500 for contributing to the study.

Participants

Study participants were identified through record reviews at four private psychological clinics in a large city in the western United States. Records for clients evaluated across the 12 months prior (January 1, 2018 to January 1, 2019) were reviewed. Inclusion criteria were clients who were referred for and received an ASD evaluation by a licensed psychologist during that time frame.

Data Collection

Data were obtained via chart reviews of clients evaluated by the clinician raters in the prior 12 months. Informed consent was obtained by the raters from parents/legal guardians for the chart reviews prior to data collection. Clinician raters then completed the CADE survey using the information from each client's chart and entered the de-identified data into an Excel spreadsheet. In addition to providing ratings on CADE items, raters provided client age, gender, diagnostic information, ADOS score, and Full Scale IQ score, as well as results from any other tests administered in the course of the original evaluation. Raters were also asked to consider the ASD symptom levels using a 10-point scale with 1 indicating low severity and 10 indicating high severity. Data were saved onto a thumb drive and collected in person from the raters to further protect confidentiality of the data.

Clients whose data were used in the present study ranged in age from 1 to 44 years (Mdn $= 8.0$, $X = 9.29$, $SD = 5.97$). The majority were male ($n = 113$; 59.2%) and most were White ($n = 144$; 75.4%). Sample demographics are presented in Table 1.

Measures

Demographics

From the client's chart, raters provided client age, and gender, diagnostic information, and ADOS scores. The Full Scale IQ score (FSIQ) was obtained from an

Table 1 Sample Demographics n and Percentages (N = 191)

Variable	N (percentage)
Gender	
Male	113 (59.2%)
Female	78 (40.8%)
Autism Spectrum Disorder (ASD) Diagnosis	
Yes	138 (72.3%)
No	53 (27.7%)*
Ethnicity	
White	144 (75.4%)
Hispanic	31 (16.2%)
African American/Black	12 (6.3%)
Asian	3 (1.6%)
Other	1 (0.5%)

Forty-one participants without an ASD diagnosis were diagnosed with other disorders including anxiety, attention deficit hyperactivity disorder, and intellectual disability; the other 12 had no diagnosis

assessment of cognitive ability (e.g., WPPSI-IV, WISC-V) conducted during the client's evaluation and was included in the analyses.

CLEAR Autism Diagnostic Evaluation (CADE)

The CADE is a 33-item measure designed for use with individuals ages 24 months to adult that assesses the core symptoms of ASD. The CADE is completed by primary caregivers, as well as clinicians. Each CADE item is rated on a 4-point Likert scale ranging from 0-Not Concerned to 3-Very Concerned. "I don't know" is also a response option. The possible score range for the ASD Total score is 0–99, with higher scores associated with more significant concerns and impairment. Example CADE items that assess social communication and interactions with others include *Struggles to use eye contact to communicate* and *Does not participate in cooperative play (avoids others in play and prefers to play alone or had concerns before the age of 6?).* Sample CADE items repetitive or restricted behaviors and interests include *Shows a restricted range of interests, likes, or showing preference for highly specific topics* and *Uses repetitive language (same phrases or words over and over).* Currently, there is no validity or reliability evidence for the CADE; thus the need for the current study.

Autism Diagnostic Observation Schedule-Second Edition (ADOS-2)

The ADOS-2 is a semi-structured measure that clinicians use to assess a child's ability to play and communicate through naturalistic observation with the use of defined sets and activities and assessment criteria (Lord et al., 2012). The ADOS comparison score (Comp) ranges from 1–10, where 1 indicates minimal-to-no evidence

of autism-related symptoms and 10 indicates a high level of impairment. ADOS Social-Communication (Soc-Com) and Restricted and Repetitive Behavior (RRB) scores range from 0 to 3 with 0 indicating "no abnormality of type specified" and 3 indicating "moderate to severe abnormality."

According to the ADOS-2 manual (Lord et al., 2012), the intraclass correlation (ICC) was used to assess the interrater reliability of the domain totals for this sample. ICC of the social affect (SA) domain was 0.97 for module 1 ($n = 63$), 0.98 for module 2 ($n = 50$) and 0.92 ($n = 66$) for module 3. The repetitive and restricted (RRB) domain had ICCs of 0.79, 0.80 and 0.91 for modules 1, 2 and 3, while the ICCs of the overall totals were 0.97, 0.96 and 0.94, respectively. The interrater reliability of classification was reported in PA for modules 1–3 from still another subsample of the same dataset: 95% for module 1 ($n = 46$ autism; $n = 13$ non-spectrum), 98% for module 2 ($n = 28$ autism; $n = 6$ non-spectrum) and 92% for module 3($n = 46$ autism; $n = 1$ autism spectrum). The ADOS-2 Comparison, Soc-Com, and RRB scores were used in the present study.

Analyses

There were eight overarching research questions for this study. To answer the first research question, "Does the CADE demonstrate adequate internal consistency?" a Cronbach's alpha was computed. Assumptions for Cronbach's alpha of items being linearly related to one another and the residuals are uncorrelated were tested. To answer the questions, "Do CADE scores differ based on gender or race/ethnicity?"; "Do CADE Total scores differ between individuals with a diagnosis of autism and those who do not have a diagnosis"; and "Do CADE Total scores differ between individuals with a Full Scale IQ score in the Intellectual Disability range (< 70) and those with Full Scale IQs above 70?", independent t tests were conducted. Assumptions of independent t tests were checked including: if the variances of the two samples were equal, normal distribution of the dependent variable, and the data being independent.

Pearson correlations were conducted to answer, "Are CADE Total scores significantly associated with ADOS Comp, ADOS Soc-Com, RRB scores?" and "Are CADE Total scores associated with Full Scale IQ scores?" Assumptions for Pearson correlation including that the two variables have a linear relationship, normally distributed scores, and no outliers were checked.

To answer the research question, "What is the ASD diagnostic accuracy of the CADE Total score?", sensitivity and specificity and area under the receiver operating characteristic (ROC) curve were calculated. ROC is widely used to evaluate diagnostic tests with a dichotomous outcome (Ma et al., 2015). ROC represents the plot of sensitivities (true positives) versus specificities (1-true negatives), the resulting area under the curve (AUC) measures test accuracy and discriminatory power. An AUC of 1 represents perfect classification while an AUC of 0.5 indicates no discriminative value (Fischer et al., 2015). Sensitivity and specificity for specific scores on the measure can be examined to determine an appropriate clinical cut-off. AUC,

sensitivities, and specificities with 95% confidence intervals (CI) were calculated for CADE total scores.

To answer the final research question of "Do CADE Total scores differ between individuals with differing levels of ASD as determined by diagnosing clinicians?" assumptions for ANOVA were checked and the variables were not normally distributed, thus, a nonparametric Kruskal–Wallis test was conducted. Due to the Kruskal–Wallis test being nonparametric, there were no assumptions to assess.

Results

A Cronbach's alpha was calculated to examine the CADE's internal consistency. The alpha was 0.94, which indicates that the items form a scale that has high internal consistency.

Independent t tests were conducted to determine if CADE Total scores differ based on gender or race/ethnicity. No statistically significant difference was found for gender or race/ethnicity (as measured by White vs. not White) in regard to CADE Total scores. Thus, males and females had similar Total scores on the CADE, and Whites and non-Whites had similar scores.

Pearson correlations assessed the relationship of CADE Total scores and ADOS scores. A statistically significantly correlation was found between CADE Total scores and ADOS Comp scores, $r(175) = 0.86$, $p < 0.001$; between the CADE Total scores and the ADOS Soc-Com scores, $r(140) = 0.70$, $p < 0.001$; and between the CADE Total scores and the ADOS RRB scores, $r(129) = 0.52$, $p < 0.001$. This means that clients who had a high score on the CADE assessment also had high scores on the ADOS assessment. Descriptive statistics for all measures are presented in Table 2. Not all measures were administered with all clients and thus sample sizes varied by measure used The r values for each of these analyses indicate a large effect size.

To assess the ASD diagnostic validity/accuracy of CADE Total scores, sensitivity and specificity and area under the ROC curve (AUC) were calculated. An AUC of 0.994 (standard error = 0.003; 95% confidence interval = 0.988–1.000) was obtained, indicating excellent discriminative capacity of the CADE total score (Streiner & Cairney, 2007). Via the ROC analysis, the optimal cutoff point for the CADE Total

Table 2 Means and Standard Deviations for Study Measures	Measure	n	M	SD	Range
	CADE Total Score	191	37.83	21.56	0–84
	ADOS Comp	175	5.89	3.15	1–10
	ADOS Soc-Com	140	10.26	5.07	1–20
	ADOS RRB	129	3.15	1.73	1–8
	FSIQ	166	97.39	23.25	30–155

ADOS Comp ADOS Comparison score, *ADOS Soc-Com* ADOS Social-Communication score, *ADOS RRB* ADOS Restricted and Repetitive Behavior score, *Wechsler FSIQ* Full Scale IQ score

ROC Curve
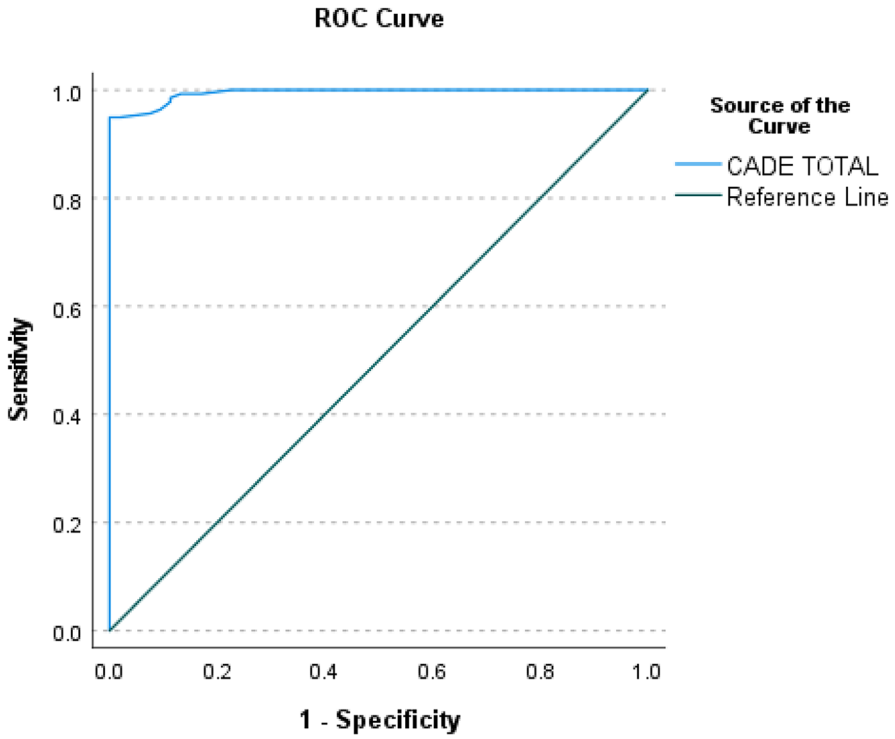
Fig. 1 Receiver operating characteristics curve for CADE Total Scores

score was identified. To discriminate individuals with ASD from those without, the optimal cutoff point was determined at a score of 25 (sensitivity = 0.949, specificity = 0.000). The ROC curve is depicted in Fig. 1.

An independent *t* test was conducted to determine if CADE Total scores differ between individuals with a diagnosis of ASD and those who do not have a diagnosis. A statistically significant difference was found, $t(176.87) = 24.58$, $p < 0.001$. This indicates that those with a diagnosis of ASD and those without a diagnosis differ on their CADE Total scores. The effect size of Cohen's *d* was 12.83, which is a large effect size.

An independent *t* test examined differences in CADE Total scores between individuals with a Full Scale IQ score in the Intellectual Disability (ID) range (< 70) and those with Full Scale IQs above 70. A statistically significant difference was found between these two groups, $t(32.88) = 4.38$, $p < 0.001$. Results indicate that those with a Full Scale IQ score in the ID range had significantly higher CADE Total scores than those who had Full Scale IQ scores above 70. The effect size of Cohen's *d* was 20.59, which is a large effect size.

To assess whether CADE Total scores differ between individuals with differing levels of ASD as determined by the diagnosing clinician, a nonparametric Kruskal–Wallis test was conducted. The test indicated that the four levels of ASD (0 = no concern and 3 = very concerned) differed on the CADE Total score, χ^2 (3,

Level of Autism	Mean	*n*	*z*	*r*
0	24.50	48	7.81*	.85
3	66.50	36		
0	24.50	48	9.09*	.85
2	66.50	36		
0	26.65	48	7.26*	.77
1	66.49	41		
1	22.73	41	6.81*	.78
3	57.53	36		
1	37.98	41	4.21*	.41
2	63.95	66		
2	38.48	66	6.02*	.60
3	75.36	36		

Table 3 Post hoc Results for Autism Level and CADE Scores

* *p* < .001

$N = 191) = 138.57$, $p < 0.001$. Six post hoc Mann–Whitney tests were conducted to assess which groups differed from one another using the Bonferroni correction with an alpha value of 0.008. Table 3 presents the post hoc test results.

Pearson correlations were computed to examine the association of CADE Total scores and Wechsler Scale Full Scale IQ scores. Assumptions were checked and met. A statistically significantly negative correlation was found between CADE total scores and Full Scale IQ scores, $r(166) = -0.37$, $p < 0.001$: participants who had a high score on the CADE assessment ($M = 37.83$, $SD = 21.56$) had a lower Full Scale IQ score ($M = 97.39$, $SD = 23.25$), and vice versa. The result is considered a medium sized effect size.

Discussion

Recent surveillance efforts indicate that 1 in 54 American children meet the criteria for Autism Spectrum Disorder (ASD), making it the fastest growing neurodevelopmental disorder in the U.S. (Maenner et al., 2020). The ASD diagnostic evaluation process can be lengthy and stressful for families. Given that intervention is generally dependent on a diagnosis of ASD (Bent et al., 2017; Hyman et al., 2020) and early identification and intensive early intervention are associated with more favorable outcomes for children with ASD (Fuller & Kaiser, 2019; Rogers & Vismara; 2008), delays in the diagnosis of ASD are of concern. The CLEAR Autism Diagnostic Evaluation (CADE; Willard, & Kroncke, 2019) was developed in response to the need to improve, shorten, and standardize the clinical ASD evaluation process. The aim of the present study was to gather initial psychometric support for the use of the CADE as a reliable and accurate measure of ASD.

The results of the present study provide preliminary evidence of technical adequacy of the CADE. The CADE has an internal consistency of 0.94, which yields a high level of acceptability based on the standard of 0.80 (Miller et al., 2013). This

result suggests that the CADE is able to consistently measure the behaviors associated with ASD.

Results provide support for the diagnostic accuracy as the CADE total score showed an excellent ROC AUC of > 0.90 for discriminating those with and without ASD. Results also provide evidence of the validity of the CADE through analysis of CADE scores in relation to other constructs involved in ASD assessment. The present study found the CADE to be well correlated with ADOS-2 scores. Given that the ADOS "has become a first-choice diagnostic instrument, often designated as 'gold standard'" (Zander et al., 2016, p. 770), this is a key finding supporting the use of the CADE in the identification of ASD. The CADE demonstrated good discriminant validity in that CADE Total scores differentiated between individuals who had a diagnosis of ASD and those who did not.

Additionally, CADE total scores were found to be significantly negatively associated with Full Scale IQ scores. This finding is consistent with research examining the relationship between ADOS scores and scores on cognitive ability assessments (e.g., Joseph et al., 2002). Similarly, the CADE Total scores differed significantly for individuals who had FSIQ scores in the Intellectual Disability (ID) range from those whose scores were ≥ 70 or above. ID is a commonly comorbid condition of ASD (Charman et al., 2011; Dykens & Lense, 2011), and the needs of individuals who have co-occurring cognitive impairment will differ and may be more intensive as they tend to demonstrate least improvement in social communication over time (Fountain et al., 2012). The results of the present study suggest that the CADE scores align with these clinical findings.

Finally, research has demonstrated that ASD symptom severity tends to decrease over time (e.g., Clark et al., 2017). CADE total scores were negatively associated with age $r(191) = -0.300$, $p < 0.001$, a finding consistent with previous research. This provides additional support for the use of the CADE in the assessment and identification of ASD. Taken together, the present study provides promising evidence for the validity of the CADE.

Limitations

There are several limitations to the findings of this study that should be considered. First, the CADE was designed to be completed by the parent/primary caregiver and then the clinician completes their own CADE considering this information, as well as their own clinical impressions as part of the evaluation. In the current study, psychologists who had previously evaluated the individual for ASD completed the CADE ratings based on a review of their own clients' files. Being privy to the results of the clients' comprehensive evaluation may have influenced their ratings; as such, future research should include administration of the CADE by a separate clinician at the same time the ADOS was administered and scored by an evaluating clinician. Additional studies are also needed to examine the convergent validity with other measures using ratings from individuals' primary caregivers. Second, all clients and raters were from the state of Colorado, limiting the generalizability of the study findings to other geographic regions. Third, more than three quarters of the clients

who were rated identified as White. Additionally, these clients received these evaluations in private clinical settings. Future research with more diverse populations and in a variety of settings is warranted. Finally, only clinicians and ASD experts were included in the CADE development process. Including other stakeholders such as parents and teachers in any refinement of the CADE is recommended.

Implications and Future Directions

The CADE appears to have practical value for the identification and diagnosis of individuals with ASD in a variety of settings. A quick and valid measure of ASD like the CADE offer an efficient means of identifying of children with ASD. Importantly, this could lead to the provision of early intervention services that can affect a child's development and outcomes. Given the evidence for the impact of early diagnosis and intervention on outcomes for individuals with ASD, it is critical to have a reliable and efficient means for identifying ASD in both clinical and school settings (Fuller & Kaiser, 2019; Rogers & Vismara, 2008).

Although a measure such as the ADOS-2 is highly regarded for diagnosis, its use requires specially trained clinicians that are expensive to maintain in a clinic or hospital. In order to have obtained essential competence to administer the ADOS-2, individuals must have Level C degree qualifications, complete either video lessons or an in-person workshop costing several hundred dollars, and obtain continuing education contact hours (Lord et al., 2012; WPS, 2018). Moreover, although the ADOS-2 demonstrates high inter-rater reliability in research settings where evaluators are provided with thorough preparation, continuous, and systematic calibration there is some evidence that the inter-rater agreement in ordinary daily clinical use of the ADOS-2 is low and highly variable (Kamp-Becker et al., 2018), suggesting diagnostic accuracy depends on specialized experience in the diagnostic evaluation. Clearly, missed or inaccurate identification might delay intervention and/or lead to the provision of ineffective services or services provided at the wrong intensity, which may significantly affect a child's developmental trajectory and academic outcomes (Granpeesheh et al., 2009). The CADE demonstrated excellent specificity and sensitivity and its use does not require intensive training and supervision and thus lends itself to a more efficient accurate diagnosis of ASD.

Additionally, implementing the ADOS-2 is often not practical for school-based assessment. Conducting the ADOS-2 requires structured observation in contrived scenarios, meaning students assessed using the ADOS-2 must be pulled from class instruction for the evaluation. On average, administration time takes at least 40 to 60 min (Lord et al., 2012). These factors may explain why most school psychologists do not engage in recommended ASD assessment practices, with most relying on ASD checklists that provide limited information (Aiello et al., 2017). With approximately 762,000 children ages 3 to 21 receiving services under the special education classification of autism during the 2018–2019 academic year (Institute of Education Sciences, National Center for Education Statistics, 2020), a comprehensive evaluation is critical to the development of a culturally

responsive, intensive individualized intervention plan. Having a user-friendly tool with demonstrated technical adequacy such as the CADE that can be used in the school setting may foster both earlier identification and effective intervention.

Another strength of the CADE is that it can be completed via telehealth, something for which the ADOS is not designed. In addition to increased opportunities to assess and diagnose ASD during the COVID-19 pandemic, reliable and accurate assessment of ASD via telehealth platforms may increase accessibility for traditionally underserved populations such as racial and ethnic minorities (Mandell et al., 2009), those in rural areas (Antezana et al., 2017), and individuals who are medically fragile and incapable of visiting a provider's office. There is preliminary evidence that telemedicine-based diagnostic consultation and assessment may accurately diagnose ASD in children (Alfuraydan et al., 2020; Juárez et al., 2018). Thus, the CADE could allow for an accurate and reliable ASD assessment for individuals from toddler to adult age from the comfort of their own homes.

The current study indicates that the CADE has practical value in identifying individuals with ASD. While this study provides evidence that the CADE is a psychometrically sound instrument, additional studies of reliability are needed. Future studies on the CADE should address limitations in the current study in order to provide additional evidence for the validity of the instrument. This should include assessing inter-rater reliability of the instrument, as well as the validity of the CADE when completed by teachers and primary caregivers. Specifically, the convergent validity of the CADE should be examined with respect to scores obtained by different raters. It is imperative that future studies also examine the generalizability of results in the present study to different geographic regions, ethnicities, and treatment settings.

Data Availability Due to privacy issues, the data used in this study are not available.

Declarations

Ethical Approval All procedures performed in studies involving human participants were in accordance with ethical standards of the Colorado Multiple Institutional Review Board in 2020 and with the 1964 Helsinki declaration and its later amendments.

Informed Consent Informed consent was obtained from all individual participants/ parents or legal guardians of the participants.

Conflict of Interest The authors have no conflicts of interest to declare.

References

Ahlers, K., Gabrielsen, T. P., Ellzey, A., Brady, A., Litchford, A., Fox, J., & Carbone, P. S. (2019). A pilot project using pediatricians as initial diagnosticians in multidisciplinary autism evaluations for young children. *Journal of Developmental & Behavioral Pediatrics, 40*(1), 1–11. https://doi.org/10.1097/DBP.0000000000000621

Aiello, R., Ruble, L., & Esler, A. (2017). National study of school psychologists' use of evidence-based assessment in autism spectrum disorder. *Journal of Applied School Psychology, 33*(1), 67–88. https://doi.org/10.1080/15377903.2016.1236307

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

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

Antezana, L., Scarpa, A., Valdespino, A., Albright, J., & Richey, J. A. (2017). Rural trends in diagnosis and services for autism spectrum disorder. *Frontiers in Psychology, 8*, 590. https://doi.org/10.3389/fpsyg.2017.00590

Baio, J., Wiggins, L., Christensen, D. L., Maenner, M. J., Daniels, J., Warren, Z., 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*, *67*(6), 1. https://doi.org/10.15585/mmwr.ss6706a1

Bisgaier, J., Levinson, D., Cutts, D. B., & Rhodes, K. V. (2011). Access to autism evaluation appointments with developmental-behavioral and neurodevelopmentalsubspecialists. *Archives of Pediatrics & Adolescent Medicine, 165*(7), 673–674. https://doi.org/10.1001/archpediatrics.2011.90

Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quiñonez, H. R., & Young, S. L. (2018). Best practices for developing and validating scales for health, social, and behavioral research: a primer. *Frontiers in Public Health, 6*, 149. https://doi.org/10.3389/fpubh.2018.00149

Bent, C. A., Barbaro, J., & Dissanayake, C. (2017). Change in autism diagnoses prior to and following the introduction of DSM- 5. *Journal of Autism and Developmental Disorders, 47*(1), 163-171. https://doi.org/10.1007/s10803-016-2942-y

Clark, M. L., Barbaro, J., & Dissanayake, C. (2017). Continuity and change in cognition and autism severity from toddlerhood to school age. *Journal of Autism and Developmental Disorders, 47*(2), 328–339. https://doi.org/10.1007/s10803-016-2954-7

Charman, T., Jones, C. R., Pickles, A., Simonoff, E., Baird, G., & Happé, F. (2011). Defining the cognitive phenotype of autism. *Brain Research, 1380*, 10–21. https://doi.org/10.1016/j.brainres.2010.10.075

Crane, L., Chester, J. W., Goddard, L., Henry, L. A., & Hill, E. (2016). Experiences of autism diagnosis: A survey of over 1000 parents in the United Kingdom. *Autism*, *20*(2), 153–162. https://doi.org/10.1177/1362361315573636

Daniels, A. M., & Mandell, D. S. (2013). Children's compliance with American Academy of Pediatrics' well-child care visit guidelines and the early detection of autism. *Journal of Autism and Developmental Disorders, 43*(12), 2844–2854. https://doi.org/10.1007/s10803-013-1831-x

Dykens, E. M., & Lense, M. (2011). Intellectual disabilities and autism spectrum disorder: A cautionary note in autism spectrum disorders. In: D. Amaral, G. Dawson, and D. Geschwind (Eds.), (pp 261–269) New York: Oxford University Press.

Fischer, A., Fischer, M., Nicholls, R. A., Lau, S., Poettgen, J., Patas, K., & Gold, S. M. (2015). Diagnostic accuracy for major depression in multiple sclerosis using self-report questionnaires. *Brain and Behavior, 5*(9), e00365. https://doi.org/10.1002/brb3.365

Fountain, C., Winter, A. S., & Bearman, P. S. (2012). Six developmental trajectories characterize children with autism. *Pediatrics, 129*(5), e1112–e1120. https://doi.org/10.1542/peds.2011-1601

Fuller, E. A., & Kaiser, A. P. (2019). The effects of early intervention on social communication outcomes for children with autism spectrum disorder: A meta-analysis. *Journal of Autism and Developmental Disorders, 50*, 1683–1700. https://doi.org/10.1007/s10803-019-03927-z

Granpeesheh, D., Dixon, D. R., Tarbox, J., Kaplan, A. M., & Wilke, A. E. (2009). The effects of age and treatment intensity on behavioral intervention outcomes for children with autism spectrum disorders. *Research in Autism Spectrum Disorders, 3*(4), 1014–1022. https://doi.org/10.1016/j.rasd.2009.06.007

Harris, S. L., & Handleman, J. S. (2000). Age and IQ at intake as predictors of placement for young children with autism: A fourto six-year follow-up. *Journal of Autism and Developmental Disorders, 30*(2), 137–142.

Hyman, S. L., Levy, S. E., & Myers, S. M. (2020). Identification, evaluation, and management of children with autism spectrum disorder. *Pediatrics*, *145*(1). https://doi.org/10.1542/peds.2019-3447

Institute of Education Sciences, National Center for Education Statistics. (2020). Children 3 to 21 years old served under Individuals with Disabilities Education Act (IDEA), Part B, by type of disability. Retrieved from https://nces.ed.gov/programs/digest/d19/tables/dt19_204.30.asp

Itzchak, E. B., & Zachor, D. A. (2011). Who benefits from early intervention in autism spectrum disorders? *Research in Autism Spectrum Disorders, 5*(1), 345–350. https://doi.org/10.1016/j.rasd.2010.04.018

Johnson, C. P., & Myers, S. M. (2007). Identification and evaluation of children with autism spectrum disorders. *Pediatrics, 120*(5), 1183–1215. https://doi.org/10.1542/peds.2007-2361

Joseph, R. M., Tager-Flusberg, H., & Lord, C. (2002). Cognitive profiles and social-communicative functioning in children with autism spectrum disorder. *Journal of Child Psychology and Psychiatry, 43*(6), 807–821. https://doi.org/10.1111/1469-7610.00092

Juárez, A. P., Weitlauf, A. S., Nicholson, A., Pasternak, A., Broderick, N., Hine, J., & Warren, Z. (2018). Early identification of ASD through telemedicine: Potential value for underserved populations. *Journal of Autism and Developmental Disorders, 48*(8), 2601–2610. https://doi.org/10.1007/s10803-018-3524-y

Kamp-Becker, I., Albertowski, K., Becker, J., Ghahreman, M., Langmann, A., Mingebach, T., & Stehr, T. (2018). Diagnostic accuracy of the ADOS and ADOS-2 in clinical practice. *European Child & Adolescent Psychiatry, 27*(9), 1193–1207. https://doi.org/10.1007/s00787-018-1143-y

Kroncke, A., Willard, M., & Huckabee, H. (2016). *Assessment of autism spectrum disorder: Critical issues in clinical, forensic, and school settings*. Springer Publishing.

Lappé, M., Lau, L., Dudovitz, R. N., Nelson, B. B., Karp, E. A., & Kuo, A. A. (2018). The diagnostic odyssey of autism spectrum disorder. *Pediatrics, 141*(Supplement 4), S272–S279.

Lord, C., Risi, S., DiLavore, P. S., Shulman, C., Thurm, A., & Pickles, A. (2006). Autism from 2 to 9 years of age. *Archives of General Psychiatry, 63*(6), 694–701. https://doi.org/10.1001/archpsyc.63.6.694

Lord, C., Risi, S., Lambrecht, L., Cook, E. H., Leventhal, B. L., DiLavore, P. C., & Rutter, M. (2000). The Autism Diagnostic Observation Schedule—Generic: A standard measure of social and communication deficits associated with the spectrum of autism. *Journal of Autism and Developmental Disorders, 30*(3), 205–223. https://doi.org/10.1023/A:1005592401947

Lord, C., Rutter, M., Dilavore, P. C., Risi, S., Gotham, K., Bishop, S. L. (2012.) Autism Diagnostic Observation Schedule, Second Edition (ADOS-2) manual (part 1) modules 1–4. *Western Psychological Services*.

Ma, H., Bandos, A. I., Rockette, H. E., & Gur, D. (2015). On use of partial area under the ROC curve for evaluation of diagnostic performance. *Statistics in Medicine, 32*(20), 3449–3458. https://doi.org/10.1002/sim.5777

Maenner, M. J., Shaw, K. A., Baio, J, Washington, A. Patrick, M., DiRienzo, M., Christensen, D.L., Wiggins, L. D., Pettygrove, S., Andrews, J. G., Lopez, M., Hudson, A., Baroud, T., Schwenk, Y., White, T., Rosenberg, C. R., Lee, L., Harrington, R. A., Huston, M., Dietz, P. M. (2020). Prevalence of autism spectrum disorder among children aged 8 years — Autism and Developmental Disabilities Monitoring Network, 11 sites, United States, 2016. *MMWR Surveillance Summaries, 69*(4), 1–12. https://doi.org/10.15585/mmwr.ss6904a1

Mandell, D. S., Wiggins, L. D., Carpenter, L. A., Daniels, J., DiGuiseppi, C., Durkin, M. S., & Shattuck, P. T. (2009). Racial/ethnic disparities in the identification of children with autism spectrum disorders. *American Journal of Public Health, 99*(3), 493–498. https://doi.org/10.2105/AJPH.2007.131243

McMorris, C., Cox, E., Hudson, M., Liu, X., & Bebko, J. N. (2013). The diagnostic process of children with autism spectrum disorder: Implications for early identification and intervention. *Journal on Developmental Disabilities, 19*(2), 42–49.

Miller, M. D., Linn, R. L., & Gronlund, N. E. (2013). Measurement and assessment in teaching (11th ed.). Pearson.

Risi, S., Lord, C., Gotham, K., Corsello, C., Chrysler, C., Szatmari, P., ... & 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

Rogers, S. J., & Vismara, L. A. (2008). Evidence-based comprehensive treatments for early autism. *Journal of Clinical Child & Adolescent Psychology, 37*(1), 8–38. https://doi.org/10.1080/153744107018178081

Siklos, S., & Kerns, K. A. (2007). Assessing the diagnostic experiences of a small sample of parents of children with autism spectrum disorders. *Research in Developmental Disabilities, 28*(1), 9–22. https://doi.org/10.1016/j.ridd.2005.09.003

Streiner, D. L., & Cairney, J. (2007). What's under the ROC? An introduction to Receiver Operating Characteristics Curves. *The Canadian Journal of Psychiatry, 52*, 121–128.

Wiggins, L. D., Baio, J. O. N., & Rice, C. (2006). Examination of the time between first evaluation and first autism spectrum diagnosis in a population-based sample. *Journal of Developmental & Behavioral Pediatrics, 27*(2), S79–S87.

Willard, M., Kroncke, A. P., & Emily L. R. Harrison, E. L. R. (2019a). Providing a CLEAR answer: Testing a novel scale to assess the accuracy of prediction of autism spectrum disorder diagnosis. Unpublished Manuscript. https://tinyurl.com/4fym6b2b

Willard, M. & Kroncke, A. (2019). Clear Autism Diagnostic Evaluation (CADE) manual. *Clear Child Psychology*.

WPS. (2018). FAQ ADOS-2 requirements. Retrieved February 12, 2021, from https://www.wpspublish.com/faq-ados-2-requirements.

Zander, E., Willfors, C., Berggren, S., Choque-Olsson, N., Coco, C., Elmund, A., ... & Bölte, S. (2016). The objectivity of the Autism Diagnostic Observation Schedule (ADOS) in naturalistic clinical settings. *European Child & Adolescent Psychiatry, 25*(7), 769–780. https://doi.org/10.1007/s00787-015-0793-2

Zwaigenbaum, L., Bryson, S., & Garon, N. (2013). Early identification of autism spectrum disorders. *Behavioural Brain Research, 251*, 133–146. https://doi.org/10.1016/j.bbr.2013.04.004

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.