



An Evaluation of Parent and Teacher Discrepancies on an Adaptive Behavior Measure for Youth with Autism Spectrum Disorder, Intellectual Disability, and Global Developmental Delay

Mallory A. Stevens¹ · Kimberly J. Selders² · Olivia Jeckel³ · Valerie Brownfield⁴ · Kerri P. Nowell⁵

Accepted: 23 March 2022 / Published online: 21 April 2022

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

Abstract

Comprehensive evaluations include data from multiple informants, but discrepancies occur on adaptive skill measures of youth with autism spectrum disorder (ASD) and intellectual disabilities (ID). This study investigated discrepancies between parent-teacher ratings on a measure of adaptive behavior [*Adaptive Behavior Assessment System, Third Edition (ABAS-3)*] in a clinical sample of 115 youth. Agreement between informants was determined and then discrepancies were identified using paired-sample *t*-tests for the whole sample and subsamples. Factors associated with parent-teacher discrepancies were investigated including age, diagnoses, IQ, autism symptomology, and parent education. Parent-teacher scores were moderately correlated for the ABAS-3 composite and domains. Teachers rated youth with ASD and ID as having higher adaptive skills. Autism symptomology significantly predicted discrepancies between informants on the ABAS-3.

Keywords Autism spectrum disorder · Intellectual disability · Adaptive behavior · Informant discrepancies

Introduction

Neurodevelopmental disorders consist of a group of disorders that occur early in key developmental periods (i.e., early childhood), impact life-long developmental processes (American Psychological Association [APA], 2013), and include autism spectrum disorder (ASD), intellectual disability (ID), and global developmental delay (GDD). ASD is characterized as deficits in social communication and interaction and the presence of restricted repetitive behaviors,

and is estimated to occur in 1 in every 54 children in the United States (Manner et al., 2020). ASD is a heterogeneous disorder and autistic youth differ in symptom presentation and other related behaviors and skills such as cognitive ability and adaptive behavior (i.e., broad conceptual, social, and practical skills required to function in everyday life). For example, 31% of children with ASD are diagnosed with ID (Baio et al., 2018) and 5.3% are diagnosed with GDD (Zablotsky et al., 2015). The identification of ID requires deficits in cognitive abilities *and* deficits in adaptive behavior that are apparent from an early age (i.e., adaptive behavior deficits define the disorder), whereas GDD is applied to youth under the age of five who demonstrate delays in multiple cognitive domains and often have corresponding deficits in adaptive functioning (APA, 2013). Although specific diagnostic criteria for ASD, ID, and GDD differ, all diagnoses require that there be evidence of functional impairment as a result of symptom presentation; thus, adaptive behavior deficits are likely to be evident across all disorders.

Comprehensive Evaluations and Adaptive Behavior

Comprehensive evaluations of ASD, ID, and GDD require professionals to assess developmental domains using data from multiple sources to make diagnostic decisions and

✉ Mallory A. Stevens
mashx5@umsystem.edu

¹ Department of Educational, School, and Counseling Psychology, University of Missouri, 16 Hill Hall, Columbia, MO 65211, USA

² Department of Special Education, University of Missouri, Columbia, MO, USA

³ Department of Psychological Services, University of Missouri, Columbia, MO, USA

⁴ Department of Public Health, University of Missouri, Columbia, MO, USA

⁵ Department of Health Psychology, Thompson Center for Autism and Neurodevelopmental Disorder, University of Missouri, Columbia, MO, USA

capture an individual's present level of functioning (Hyman et al., 2020; Jordan et al., 2019; McDonald et al., 2016; Wilkinson, 2017). Youth with these suspected neurodevelopmental disorders (i.e., ASD, ID, GDD) may present with developmental differences across a wide range of skills. As such, best practice evaluations include comprehensive assessment across myriad domains including developmental history, cognition, language, academic achievement, social, emotional, and behavioral functioning, and adaptive behavior skills (Wilkinson, 2017). Obtaining reports on standardized measures about client behaviors and skills from informants who know the individual is a critical aspect of best practice evaluations for ASD, ID and GDD, and this is especially important regarding adaptive behavior skills (Floyd et al., 2015; Jordan et al., 2019; Stratis & Lecavalier, 2015).

Adaptive behavior refers broadly to three primary factors (i.e., conceptual, social, and practical skills) required to function in everyday life (Schalock et al., 2010). Conceptual skills include language (expressive and receptive), academic skills such as reading and writing, telling time, and understanding numbers. Social skills not only encompass interpersonal skills (e.g., going places with friends) but also safety related skills such as understanding when one is in an unsafe social situation or when one is not welcome in a social group. Practical skills include other common activities of daily living such as toileting, dressing, cooking, and following schedules. As noted, clinical data from measures of adaptive behavior are necessary when making differential diagnostic decisions (e.g., documented deficits in adaptive behavior is a requirement for a diagnosis of ID), but also essential in developing recommendations for subsequent intervention and ensuring that supports are available in both home and school settings as appropriate (Floyd et al., 2015; Jordan et al., 2019). Deficits in adaptive behavior are not an operational requirement for an ASD diagnosis (i.e., the *Diagnostic and Statistical Manual for Mental Disorders-5* does not require deficits in adaptive behavior be documented using specific measurement); however, adaptive behavior difficulties are pervasive in ASD. Clients with ASD and average cognitive abilities have also demonstrated difficulty in adaptive behavior (Kanne et al., 2011). In other words, despite meeting benchmarks in cognitive abilities, autistic individuals have unexpected deficits in adaptive skills and still struggle with the basic skills needed to get through an entire day.

Deficits in adaptive behavior impede daily living skills (e.g., bathing, feeding, shopping) as well as attainment of independence and autonomy. Adaptive behavior data offers information about how an individual's differences and diagnosis specifically impact their daily living and functioning. Data from measures of adaptive behavior can be used to contribute to the clients' strengths and weaknesses profile

and can help to determine long term goals and a prognosis on their ability of being independent (Saulnier & Klaiman, 2018). Specific skills can be targeted to increase the likelihood of independent living in the future, and it is not uncommon for adaptive skills to be prioritized when planning intervention.

Measurement of Adaptive Behavior

Without reliable data, an individual risks not receiving an accurate diagnosis and the targeted intervention needed to optimize their long-term outcomes. Adaptive behavior should be assessed using standardized measures—though observation is also often included in assessment—completed by individuals who have adequate knowledge of an individual's functioning in various settings (Tasse et al., 2012). Ideally, data is collected from multiple raters because particular skills may be more salient in different settings (e.g., dressing is required in the home but may not be at school). As such, standardized rating forms may be completed by parents, teachers, or by the individual themselves (i.e., self-report). However, when multiple informants are involved, there may be differences in reporting, and, as a result, challenges of interpreting and reconciling discrepant data. Such discrepancies are possible whenever multiple informants are used and are not only found when measuring adaptive behaviors. For example, in a review of 49 studies examining the agreement between informants rating youth with autism emotional and behavioral concerns, there were significant differences across similar (parent/parent), different (e.g., teacher/parent), and all informants for externalizing behavior ($r=0.39-0.64$), internalizing behavior ($r=0.31-0.66$), social skills ($r=0.28-0.42$), and total behavior ($r=0.36-0.64$) (Stratis & Lecavalier, 2015). Moreover, in a study evaluating informant agreement on the *Behavior Assessment System for Children 2nd Edition* (BASC-2; Reynolds & Kamphaus, 2004), reports on internalizing and externalizing behaviors were consistent between informants, however, teachers reported their students as having higher adaptive skills (McDonald et al., 2016).

On measures specific to adaptive behaviors, such as the *Adaptive Behavior Assessment System, Third Edition* (ABAS-3; Harrison & Oakland, 2015) and *Vineland Adaptive Behavior Scales-Second Edition* (Vineland-II Sparrow et al., 2006), there is some evidence of discrepancies between parent and teacher ratings in autistic youth (Dickson, 2018; Jordan et al., 2019). Specifically, one study found that teachers rated youth with ASD as having more overall skills (i.e., higher scores on the general adaptive behavior composite) and more Practical skills, but found no differences in ratings between teachers and parents on the ABAS-3 Social and Conceptual domains (Jordan et al., 2019). While there were discrepancies in the composite

adaptive behavior and Practical scores, their inclusion of intra-class correlation coefficients (ICCs) indicated moderate reliability between informants on the ABAS-3 composite, Conceptual, Social, and Practical scores. Dickson (2018) found that parent and teacher ratings were significantly different at two separate time points on the Vineland-II, with parents consistently reporting more skills than teachers. The previous studies were limited to samples of autistic youth with IQ scores higher than 70, meaning that these samples excluded youth with ID. However, there is some evidence that a pattern of higher scores on teacher ratings of adaptive behavior is evident in youth with lower cognitive abilities. For example, Ellison (2016) found that teachers rated 67 students with autism (IQ range = 38–128) as having more adaptive skills compared to their parents' ratings ($d = 0.42\text{--}0.78$) on the BASC-2.

As described above, there is evidence of multiple informant discrepancies in samples of youth with ASD, but whether diagnoses and other individual factors influence informant discrepancies on adaptive behavior measures is largely unknown. It is critical that clinicians and other stakeholders relying on data from adaptive behavior measures understand what individual child factors might result in discrepant scores across raters so that the data is interpreted in meaningful ways. There are indications that youth characteristics may play a role in the degree of concordance between informants' ratings on several measures of behavior (Stratis & Lecavalier, 2015). Meta-analytic research suggests that a clinical diagnosis of ASD and ID moderates the degree to which pairs of informants (i.e., child, parent, teacher) agree across several measures of internalizing and total behavior measures (Stratis & Lecavalier, 2015). Specifically, there may be higher agreement on *total* scores from measures evaluating broad-band externalizing and internalizing behaviors for both autistic youth and youth with ID, as well as on scales assessing internalizing behaviors for youth with ASD (Stratis & Lecavalier, 2015). Importantly, this meta-analysis did not include measures specific to adaptive behaviors.

Data from adaptive behavior measures certainly offer more insight to a child's daily living skills, as well as provide more insight and context for many other related skills like communication, cognitive, and social skills. In some ways adaptive behavior skills are the functional outcome of these related skills, which are commonly targeted in intervention. For example, if a child is not able to communicate their wants and needs this may suggest that their speech and language functioning may be impacted. This then becomes an adaptive behavior concern when the child is unable to use verbal/nonverbal strategies to complete daily living skills (e.g., requesting a snack; telling an adult when they need to use the restroom). Moreover, engagement and maintenance of conversation is considered an important skill for both adaptive and social behavior, and example of functional

contexts include social relationships (e.g., peers at school) and employment. Unsurprisingly, both speech/language and social skills are represented on many adaptive measures. Given what we know about adaptive behavior and related skills, adaptive behavior likely varies across child variables like cognitive ability, age, and language ability and these child variables have also been identified as predictors for adaptive behavior in autistic youth with varying cognitive abilities (e.g., Liss et al., 2001). Moreover, a handful of studies have specifically examined factors that may moderate concordance between respondents' ratings on adaptive behavior measures with inconsistent results. Dickson found that child cognitive ability, as measured by the *Mullen Scales of Early Learning*; (Mullen; Mullen, 2001) and *Differential Ability Scale, Second Edition* (DAS-II; Elliot, 2007), predicted parent/teacher discrepancies on adaptive behavior (Vineland-II) and autism specific rating scales. Specifically, higher cognitive ability was related to *decreased* discrepancies in ratings of youths' adaptive behavior, restricted repetitive behaviors, and social approach/withdrawal behaviors (Dickson et al., 2018). However, results from studies using the ABAS-3 found both child (i.e., age, IQ, language level, ASD symptomology) and family factors (i.e., parent education level) *did not* predict discrepancies in ratings (Hattier et al., 2013; Jordan et al., 2019; McDonald et al., 2016).

Current Study

Information about adaptive behavior is essential for best practice clinical evaluations and intervention planning for youth with ASD, ID, and GDD. Previous research has documented informant discrepancies on measures of emotional/behavioral functioning and measures of adaptive behavior (e.g., the ABAS-3). Understanding patterns of ratings across respondents is an important component of an evaluation as clinicians are often tasked with reconciling discrepant data when they are making important decisions about diagnosis and intervention. This line of research has established that use of multi-informant ratings provides meaningful information about youth behavior and discrepancies commonly represent differences across settings and context (De Los Reyes, 2011), which directly informs clinical treatment recommendations. Data obtained from multi-informants allows for specific recommendations to support skill acquisition in the settings the child is observed (i.e., schools, home). For example, adaptive behavior measures are commonly referenced for school-based services (e.g., individualized education plan). Furthermore, depending on an individual's diagnosis, there may be particular patterns in respondent ratings that could help inform clinical decision making. However, research to date is inconsistent and generalizability of findings is limited by sample characteristics. For example, to date, study samples have primarily included youth with

ASD and higher cognitive abilities. It is important to include youth with ID, because cognitive ability has been identified as a significant predictor of informant discrepancies on rating scales (Dickson et al., 2018). Moreover, there is limited research including younger children who are identified with GDD. ASD, ID and GDD are highly comorbid (Baio et al., 2018; Zablotsky et al., 2015). For this reason, more research is needed that captures adaptive presentation in autistic youth with comorbid diagnoses (e.g., ASD and ID) and different diagnoses that overlap in symptomology (e.g., GDD).

The current study includes a well-characterized clinical sample of youth with ASD, ID, and/or GDD to extend on research evaluating informant discrepancies on adaptive behavior rating scales to help providers and caregivers better support youth with adaptive behavior deficits. We did this by replicating a comprehensive, rigorous data analysis plan (Jordan et al., 2019; McDonald et al., 2016) that goes beyond use of single metric of informant agreement (i.e., correlations) by including: the examination of intra-class correlation coefficients (ICCs), between-group comparisons, and Bland–Altman plots and regressions (Bland & Altman, 1986). The current study extends on previous research with the inclusion of linear regressions to determine whether age, diagnoses, IQ, autism severity, and/or parent education predict concordance between parent and teacher ratings on the ABAS-3. The following research questions are addressed: (1) What is the level agreement between parent and teacher ratings on the ABAS-3? (2) Are discrepancies between informants associated with the child's diagnosis? (3) What are the trends in the magnitude of discrepancies across the entire sample? and (4) Do client factors predict parent/teacher discrepancies on the ABAS-3?

Method

This cross-sectional research study was determined to be exempt and approved by the university Institutional Review Board. The research questions were addressed with data from 115 participants from a well-characterized clinical sample. Each participant was evaluated at an autism specialty clinic in the Midwest. The data was gathered for secondary analysis from an existing database housing clinical data for the specialty center. The participants, measures, and procedures are described below.

Participants

Youth who participated in diagnostic clinics at an interdisciplinary, specialty-care center in the Midwest and who received diagnoses of ASD, ID, or GDD were included in this study. Inclusion criteria also required complete parent and teacher reports on the ABAS-3 (Harrison & Oakland,

2015). The resulting sample consisted of 115 youth between 2 and 19 years of age ($M=9.01$ years old; $SD=3.53$ years). Participants were predominantly White (83%) and male (83%), and most families obtained specialized training and/or a partial college training. Most participants received an ASD diagnosis ($n=91$), followed by comorbid ID and ASD ($n=13$), GDD and ASD ($n=6$), ID diagnosis ($n=3$), and GDD diagnosis ($n=2$). See Table 1 for descriptive statistics and details of the participants. In the current study the diagnosis(es) variable was dichotomized to address a gap in the literature related to informant discrepancies and to evaluate parent-teacher ratings for youth with and without lower cognitive functioning as a complete sample and separate subsamples. Participants with only ASD (i.e., without comorbid ID or GDD) represented one group ($n=91$) and all other subsamples in the other ($n=24$).

Clinical diagnoses were established using best-practice evaluations for developmental concerns and included (at minimum) a clinical diagnostic interview and comprehensive developmental history, cognitive or developmental assessment, and adaptive behavior assessments. For those patients for whom ASD was being evaluated, an *Autism Diagnostic Observation Schedule- Second Edition* (ADOS-2; Lord et al., 2012) was also administered. In addition, clients presenting with other concerns (e.g., executive functioning, anxiety, depression) were further assessed in those areas. Clinical diagnoses were made by licensed psychologists or pediatricians practicing at the autism specialty center. Participants with only an ASD diagnosis had, on average, parent adaptive behavior composites that fell in the low average range ($M=73.88$, $n=91$) and teacher adaptive behavior composites that fell in the below average range ($M=83.23$, $n=91$). Overall, the average adaptive behavior composites for youth with ID, GDD, and comorbid diagnoses fell in the extremely low range for both parent ($M=62.25$, $n=24$) and teacher ratings ($M=65.54$, $n=24$) on the ABAS-3.

Measures

Clinical reports for clients diagnosed with ASD, ID, and/or GDD seen in the past five years (2015–2020) were reviewed for demographic data, autism screening results, and measures of intellectual/developmental, adaptive behavior ability. Parents received and completed a background history form upon their child's (i.e., participants') diagnostic evaluation, and relevant demographic data extracted for the purposes of the current study included age of the child, child's gender, child's race/ethnicity, and parent education level. For the purposes of this study, the following measures were examined: the ABAS-3, the *Social Communication Questionnaire-Lifetime* (SCQ-Lifetime; completed by the parent; Rutter et al., 2003a, 2003b), and intellectual/developmental measures administered in the participants' evaluation.

Table 1 Demographic data for the participants

Variable	n	% of total	
Sex			
Female	20		17.39
Male	95		82.61
Race/ethnicity			
White	95		82.61
Black/African American	7		6.09
Latinx/Hispanic	3		2.61
American Indian/Alaskan native	1		0.87
Asian	2		1.74
Multiple	7		6.09
Diagnosis(es)			
ASD	91		79.13
ID	3		2.61
GDD	2		1.74
ASD & ID	13		11.30
ASD & GDD	6		5.22
Parent education			
Some high school	3		2.60
High school/GED	23		20
Partial college of specialized training	38		33.04
Standard college or bachelor's degree	27		23.48
Graduate or professional training	22		19.13
	<i>M</i>	<i>SD</i>	Range
Verbal IQ	87.17	23.51	6.35–124
Nonverbal IQ	85.48	22.13	19.18–131
SCQ-Lifetime	16.64	7.11	2–34

Total sample consisted of 115 participants. Several participants were missing variables including family income ($n = 14$), verbal IQ ($n = 28$) and SCQ-lifetime ($n = 23$)

ASD autism spectrum disorder, ID intellectual disability, GDD global developmental disability, SCQ-Lifetime Social Communication Questionnaire-Lifetime

Complete General Adaptive Behavior Composite (GAC) and domain ratings and scores (i.e., Conceptual, Social, and Practical domains) for both the parent and teacher ABAS-3 forms were required to be included in the current study.

Adaptive Behavior Assessment System, Third Edition (ABAS-3)

The primary measure of interest for the current study—the ABAS-3—is a norm-referenced instrument used to assess adaptive skills that are necessary for independent functioning in daily living (Harrison & Oakland, 2015). The ABAS-3 is used to inform diagnostic and eligibility decisions during comprehensive evaluations, as well as assist in formulating intervention plans based on identified strengths and challenges. This measure spans the age range of birth to 89 years of age and consists of five different rating forms that can be used to obtain input from critical informants as relevant for

the individual's age and environments. For this study, rating forms were obtained from parents and teachers. The ABAS-3, Parent/Primary Caregiver Form was administered for participants ages 0–5 years and the ABAS-3, Parent Form was administered for participants between 5 and 21 years. In addition, the ABAS-3 Teacher/Daycare Provider Form was administered for participants between 2 and 5 years of age, and the ABAS-3 Teacher Form was administered to participants between the ages of 5 and 21 years. Respondents are asked to rate the individual's abilities on a 4-point scale (0–3), ranging from “is not able to” perform the skill up to “always or almost always” performs the skill in question.

Ratings on the ABAS-3 result in raw scores that are converted to scaled scores for 9 specific adaptive skill areas ($M = 10$, $SD = 3$), as well as standard scores for adaptive skill domains of Conceptual, Social, and Practical skills ($M = 100$, $SD = 15$). Scores are compiled to form the GAC, an indicator of overall adaptive functioning ($M = 100$, $SD = 15$). Scores

obtained on the skill areas, domains, and the GAC compare the individual's reported skills to those in the normative sample of the same age. The present study compares Parent and Teacher standard scores for the Conceptual domain, Social domain, Practical domain, and GAC. The Conceptual domain assesses communication, academic, and self-management skills. The Social domain assesses interpersonal, leisure, and relationship skills. The Practical domain assesses community functioning, safety, and self-care skills.

The ABAS-3 was standardized using a standardization sample that was representative of the U.S. population except for White individuals and people from higher educational backgrounds being overrepresented (Harrison & Oakland, 2015). In terms of reliability, across all the standardizations samples, the ABAS-3 GAC has an internal consistency between 0.96 and 0.99. The reliability coefficients for adaptive domains vary by rating form and range between 0.72 and 0.99. Test–retest reliability also shows acceptable ranges of reliability (0.62–0.86). Interrater reliability estimates of GAC and adaptive domains fall in the moderate to strong level of consistency (0.67–0.92), varying by rating form. Cross-form consistency is of relevance to the present study as it measures the consistency between different informants using different forms. Cross-form consistency between parent and teacher ratings fell in the low to moderate range (0.41–0.57), which Harrison and Oakland (2015) note is to be expected and provides evidence for the need to obtain ratings from both parents and teachers during a comprehensive evaluation.

Validity of the ABAS-3 was established, in part, through analyzing intercorrelations for each skill area, all domains, and the GAC (Harrison & Oakland, 2015). Harrison and Oakland (2015) note that the intercorrelation findings support the expected among skills and domains, while also indicating that each skill area measures a separate construct. Factor analysis revealed that a one-factor model best fit the observed data from the standardization sample, which is consistent with prior research and conceptualization of most adaptive functioning assessments measuring a global factor. However, a three-factor model was also a close fit to the standardization sample data and supports the interpretation of three separate adaptive domains. In addition, assessment of the extent to which scores on the ABAS-3 were associated with scores on other measures of similar abilities (ABAS-II and Vineland-II) showed evidence of moderate to strong convergent validity. Finally, multivariate analysis of variance (MANOVA) analyses of the ABAS-3 also showed that adaptive composite scores were able to differentiate children with delays when compared to typically developing children. Specifically, youth with neurodevelopmental disorders were rated lower on the ABAS-3 compared to neurotypical youth with large effects for subsamples of youth diagnosed with ASD ($ES = 1.48\text{--}2.51$) and ID ($ES = 2.26\text{--}2.91$). Youth

with ADHD scored lower than the neurotypical group, but significant effects were only evident for the self-direction subscale on the parent and teacher reports ($ES = 1.53, 1.71$ respectively).

Autism Symptom Severity Screener

Parents ratings of current ASD symptoms and behaviors were assessed using the SCQ-Lifetime (Rutter et al., 2003a, 2003b). The SCQ-Lifetime is based on the *Autism Diagnostic Interview-Revised* (ADI-R; Rutter et al., 2003a, 2003b) algorithm items that gather parent report of social interaction, communication, and restricted, repetitive behaviors. More specifically, the SCQ-lifetime is designed to capture an individual's developmental history and symptom presentation between the ages of 4 and 5. The authors of the SCQ-Lifetime recommend that children be referred for a comprehensive evaluation if they receive a cutoff score of at least 15.

Measures of Intellectual and Developmental Assessment

Due to the retrospective nature of the study and variability in assessment data obtained, several different measures of intellectual and developmental functioning were used across participants. These included verbal and nonverbal index scores from the *Differential Ability Scales, Second Edition* (DAS-II; Elliot, 2007) early years ($n = 16, M \text{ Verbal} = 81.88, SD \text{ Verbal} = 24.56, M \text{ Nonverbal} = 81.88, SD \text{ Nonverbal} = 24.56$) and school-age versions ($n = 3, M \text{ Verbal} = 106, SD \text{ Verbal} = 13.53, M \text{ Nonverbal} = 106, SD \text{ Nonverbal} = 7$), *Leiter International Performance Scale, Third Edition* (Leiter-3; Roid et al., 2013; $n = 28, M = 74.5, SD = 18.40$), *Wechsler Abbreviated Scale of Intelligence, Second Edition* (WASI-II; Wechsler, 2011; $n = 35, M \text{ Verbal} = 93.20, SD \text{ Verbal} = 14.29, M \text{ Perceptual Reasoning} = 94.94, SD \text{ Perceptual Reasoning} = 14.59$), *Wechsler Intelligence Scale for Children, Fifth Edition* (WISC-V; Wechsler, 2014; $n = 22, M \text{ Verbal} = 93.73, SD \text{ Verbal} = 19.16, M \text{ Visual Spatial} = 90.95, SD \text{ Visual Spatial} = 20.20$), *Wechsler Preschool and Primary Scale of Intelligence, Fourth Edition* (WPPSI-IV; Wechsler, 2012; $n = 3, M \text{ Verbal} = 87, SD \text{ Verbal} = 10.82, M \text{ Visual Spatial} = 92.33, SD \text{ Visual Spatial} = 15.28$), and *Wechsler Adult Intelligence Scale, Fourth Edition* (WAIS-IV; Wechsler, 2008; $n = 5, M \text{ Verbal} = 87.60, SD \text{ Verbal} = 20.20, M \text{ Visual Spatial} = 85.20, SD \text{ Visual Spatial} = 23.35$).

The *Mullen Scales of Early Learning*; (Mullen; Mullen, 2001) is a developmental measure that assesses young children's Expressive and Receptive language, Fine and Gross Motor, and Visual Reception abilities. The Mullen is unlike the intelligence tests listed above, as participants' performance on subtests are reported as t-scores. To account for these differences, ratio verbal and nonverbal IQs were

calculated from the Mullen Expressive language and Visual receptive subtests respectively. Ratio verbal and nonverbal IQs were calculated by taking the age equivalent in number of months, dividing it by the clients' age at the time of their evaluation and multiplying by 100. This was completed for three participants (M Expressive = 25.61, SD Expressive = 17.48; M Visual Receptive = 40.46, SD Visual Receptive = 3.44).

Procedures

Approval was granted from the university's Institutional Review Board to analyze secondary data from a large database housed within the autism specialty clinic. Data were collected within the context of clinical evaluations where the primary referral concerns included behaviors related to ASD, developmental concerns, and comprehensive neuropsychological assessments. Prior to or during these evaluations, parents were given the ABAS-3 protocol to complete. Similarly, teacher forms were sent to teachers prior to the clients' evaluation date and returned to the primary clinician. The completed ABAS-3 forms were checked for errors and scored using scoring software. Participant data was de-identified and imported to SPSS 26 for statistical analysis.

Data Analysis Plan

Descriptive statistics were computed for participant demographic data as well as the measures used in the current study. Each participant received an individualized evaluation, and the primary referral concern was variable, which resulted in missing variables across the sample. The evaluations were not identical because of clinical decision-making by psychologists at the time of the evaluation, and organization changes across the 5 year period. It is not uncommon for autism clinicians to adjust their assessment plan to match the individual need and presentation of their client. For example, a nonverbal IQ test (e.g., *Leiter International Performance Scales, Third Edition*) may be more appropriate for clients with limited verbal abilities, and a developmental test (e.g., Mullen) for younger children. As a result, verbal and nonverbal IQ standard scores were not available for every child.

The current study replicated a comprehensive analysis plan completed by Jordan et al. (2019) and McDonald et al. (2016). A power analysis was conducted to determine the statistical power required for mean score comparisons, informant correlations, and for the multiple regression. The statistical power for the mean score comparisons was estimated at 0.98 across the total sample [$d = 0.4$, $\alpha = 0.0125$, (two-tailed), and $N = 115$ pairs; (G*Power 3.1.9.6, Faul et al., 2007)]. For the correlations between pairs, the estimated statistical power was 0.92 to detect a relationship as small as 0.3. Post hoc power analyses were

conducted for the multiple regressions conducted in this study. The sample size was 115 and 2 to 4 predictors were used with an alpha level of $p < 0.5$. The post hoc analyses revealed medium to large effects for the GAC, Conceptual, and Practical score differences ($f^2 = 0.27, 0.23, \text{ and } 0.36$ respectively), indicating adequate power (0.99–1.0). However, post hoc analysis revealed a small effect for the Social domain ($f^2 = 0.10$; Cohen, 1988) and power detection of 0.87.

Agreement and consistency between parents and teachers (i.e., research question one) were evaluated with intraclass correlations coefficients (ICC). A one-way random effects model was used for the ICC because there were different informants across the entire sample. An acceptable reliability between informants on the average measures is at least 0.70 (Koo & Li, 2016). To test whether discrepancies between informants were associated with child diagnosis (i.e., research question two) we assessed parent and teacher differences were assessed with a series of analyses including a paired samples t -test of the GAC, Conceptual, Social, and Practical domains for the entire sample and a split sample based on diagnostic category (i.e., ASD or other). Correlations alone are not sufficient to measure agreement (Altman & Bland, 1983; Ranganathan et al., 2017) as correlations only refer to the direction and magnitude of the relationship between variables. ICCs and Bland–Altman plots and correlations were generated to examine the data for trends in the data by regressing the mean ABAS-3 scores on the parent-teacher difference scores. (i.e., research question three). Similar to previous research (Jordan et al., 2019; McDonald et al., 2016), the difference scores between parent and teacher pairs were plotted on the y-axis and the mean score of each pair was plotted on the x-axis. In other words, Bland–Altman plots are a statistical method to estimate agreement between two measures where the x-axis represents how well the measures agree on average, and the y-axis represents the exact discrepancy value for each parent-teacher pair (Bland & Altman, 1986). Bland–Altman plots include upper and lower limits of agreement (LOA), which shows how much difference is displayed across the entire sample based on the mean observed difference. Four Bland–Altman plots were generated to represent informant pairs' ratings on the GAC and three ABAS-3 domains. As in Jordan et al. (2019), the scatterplots and accompanying correlations were used to further evaluate whether predictor variables moderated the relationship between parent and teacher discrepancies on the ABAS-3. Finally, we tested if any client factors predicted parent/teacher discrepancies on the ABAS-3 (i.e., research question four) by building a linear regression model. Predictors were included in the model if there was a significant correlation between the predictor and magnitude of discrepancy.

Results

Parent and Teacher Reliability Estimates

Parent and teacher reliability was analyzed with ICCs between the informants’ ratings on the ABAS-3 GAC and three domains. For all scores on the ABAS-3, there were significant correlations between the parent and teacher ratings. However, poor reliability (i.e., below the acceptable reliability of 0.70) was found between parent and teacher informants on the GAC, Conceptual, Social, and Practical domains with ICCs (Koo & Li, 2016). The average ICC between parent and teacher GAC was 0.60 with a 95% confidence interval from 0.42 and 0.72 ($F [114]=2.47, p < 0.0001$). The ICC between the average measure on the three ABAS-3 domains was 0.67 (95% CI 0.50–0.76; $F [114]=2.90, p < 0.0001$) for parent/teacher Conceptual scores, 0.43 (95% CI 0.18–0.61; $F [114]=1.77, p < 0.01$) for parent/teacher Social scores, and 0.61 (95% CI 0.44–0.73; $F [114]=2.58, p < 0.0001$) for parent/teacher Practical scores.

Cross-Informant Group Comparisons

Paired sample *t*-tests were conducted to further evaluate the discrepancy between parent and teacher ratings on the ABAS-3. Cross-informant comparisons were conducted on the GAC and three domains for the entire sample, as well as subsamples of participants with only an ASD diagnosis and comorbid presentations (refer to Table 2). Teachers rated youth with neurodevelopmental disorders as having higher Conceptual, Social, Practical, and general adaptive skills (i.e., GAC) compared to their parents’ ratings, and all comparisons were significant for the total sample with small to medium effect sizes. Similarly, teachers rated the subsample of youth with only ASD as having higher levels of adaptive skills on the GAC ($d=0.70$) as well as on the Conceptual, Social, and Practical domains ($d=0.47, 0.68, 0.13$). Finally, teachers rated autistic youth with ID or GDD

as having more skills overall, but parent-teacher ratings for the Social domain was the only significant difference ($p < 0.05$) with a small effect size of $d=0.47$.

Trends in Parent-Teacher Ratings

Bland–Altman plots were created to determine whether there were systematic trends across the informants’ difference and mean scores on the GAC and three ABAS-3 domains for the total sample. The difference scores were calculated by subtracting the teacher scores from the parent. Thus, negative difference scores represent an informant pair with higher teacher ratings. Refer to Online Resources 1–4 for sample data that illustrate the parent, teacher, difference, and mean scores used for the Bland–Altman plots. The vertical (y) axis represents the informants’ difference scores; the horizontal (x) axis represents the mean score for each pair of ratings (Figs. 1, 2, 3, 4). The solid line represents the mean difference scores, the dotted lines represent the upper and lower limits of agreement (LOA) for the difference scores, and the diagonal dotted lines represent the regression line.

Visual analysis indicated large difference ranges for the GAC, Conceptual, Social, and Practical parent and teacher ratings. For example, on the GAC (Fig. 1), there are 55 points separating the upper LOA (19.66) and lower LOA (-35.85), which suggests high variability in the difference scores and less agreement between informants. This is also suggested by the magnitude of the difference scores, which supports the findings from the paired sample *t*-tests that informant scores are significantly different. The mean difference scores for all four plots are negative (i.e., teacher ratings tend to be higher). In addition to calculating the mean difference value, upper and lower LOA were calculated to create a visual representation of the difference between the parent and teacher difference scores and 95% of future measurement pairs. Anything outside the LOA would be considered unacceptable agreement (Bland & Altman, 1986). Figure 1 displays the GAC difference and mean scores between parent and teacher pairs. All the

Table 2 Parent-teacher discrepancies across ABAS-3 scores

Composite/Domain	Total Sample					ASD subsample					Other subsample (co-occurring ID or GDD)				
	Parents		Teachers		<i>t</i> (114)	Parents		Teachers		<i>t</i> (90)	Parents		Teachers		<i>t</i> (23)
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
GAC	71.45	12.31	79.54	16.57	-6.13*	73.88	11.02	83.23	15.40	-6.28*	62.25	12.83	65.54	13.23	-1.52
Conceptual	72.50	12.20	77.86	15.82	-4.23*	75.16	11.15	81.45	15.06	-4.11*	62.38	10.79	64.25	10.48	-1.14
Social	73.03	11.27	81.40	15.93	-5.90*	74.84	10.70	83.87	15.55	-5.32*	66.17	10.93	72.04	14.04	-2.68*
Practical	74.63	14.01	83.01	18.30	-5.72*	77.08	12.63	87.35	16.53	-6.19*	65.33	15.32	66.54	15.29	-0.45

Standards for Cohen’s *d*: small = 0.20, medium = 0.50, large = 0.80 (Cohen, 1988)

* $p < 0.5$

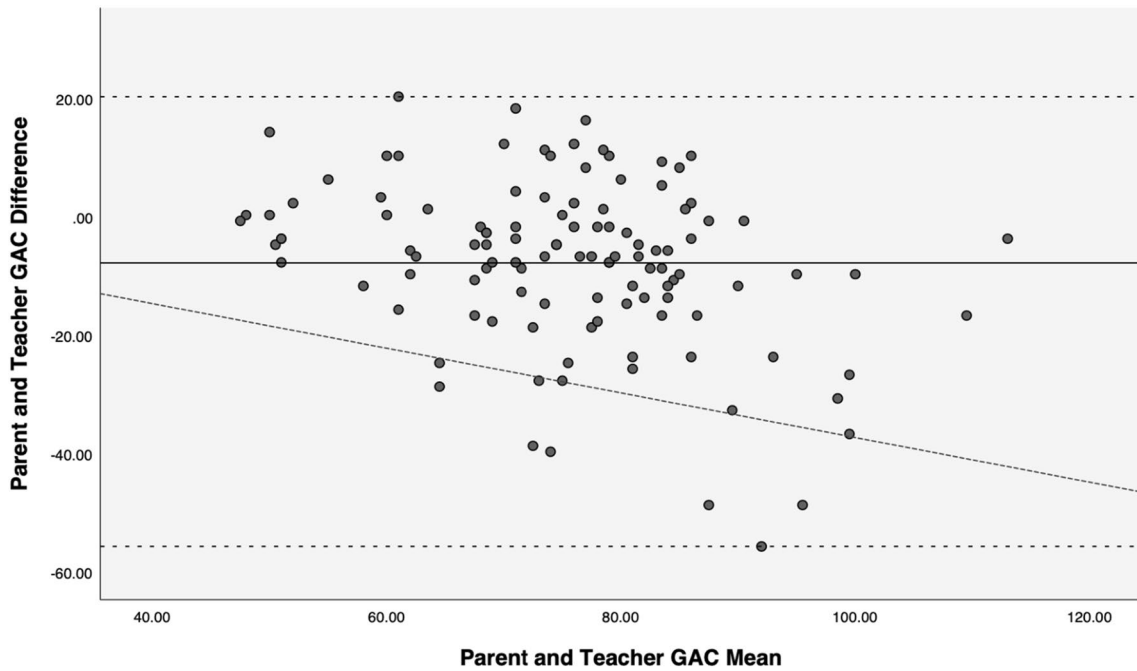


Fig. 1 Bland–Altman plot of parent/caregiver-teacher scores on the General Adaptive Composite. Scatter plot with a light gray background and dark gray data points representing parent-teacher GAC difference scores regressed onto the mean scores. The plot includes

one solid black line representing the mean difference score, two dotted lines representing the limits of agreement (or confidence intervals), and one diagonal, dotted line representing the regression equation

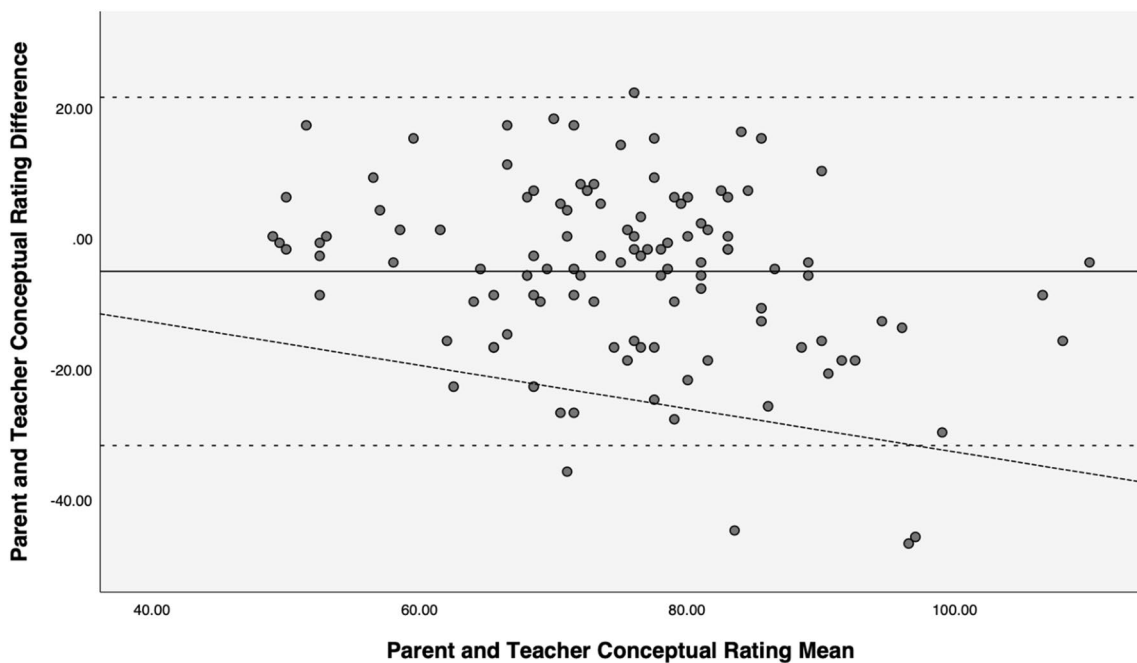


Fig. 2 Bland–Altman plot of parent/caregiver-teacher scores on the ABAS-3 conceptual domain. Scatter plot with a light gray background and dark gray data points representing parent-teacher conceptual difference scores regressed onto the mean scores. The plot

includes one solid black line representing the mean difference score, two dotted lines representing the limits of agreement (or confidence intervals), and one diagonal, dotted line representing the regression equation

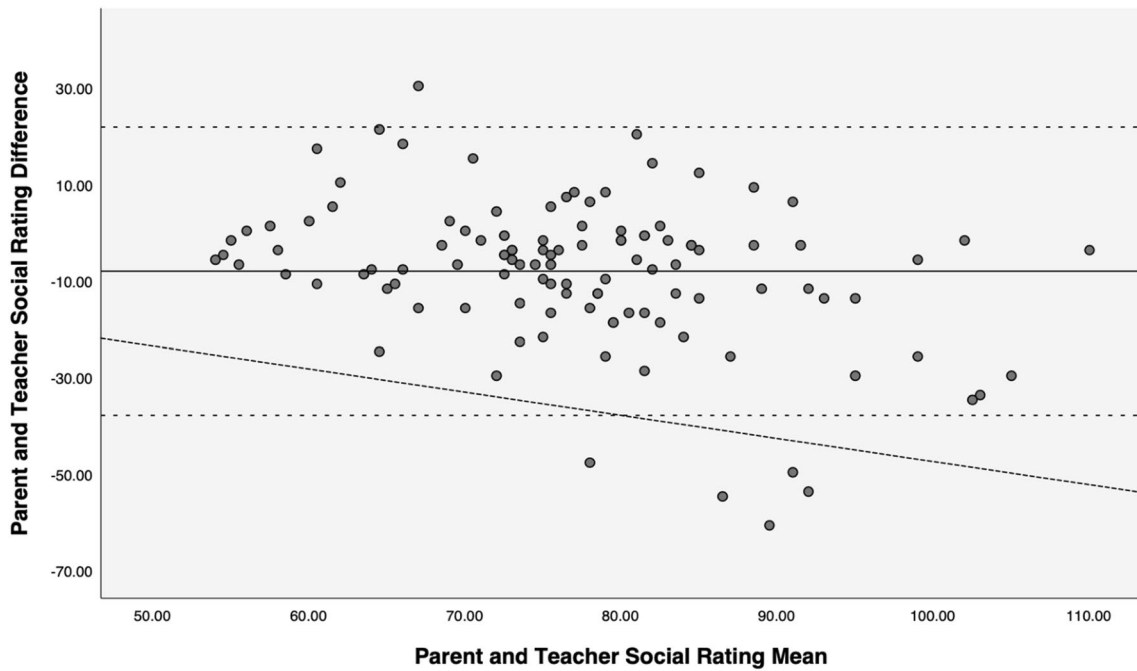


Fig. 3 Bland–Altman plot of parent/caregiver-teacher scores on the ABAS-3 social domain. Scatter plot with a light gray background and dark gray data points representing parent-teacher Social difference scores regressed onto the mean scores. The plot includes one

black line representing the mean difference score, two dotted lines representing the limits of agreement (or confidence intervals), and one diagonal, dotted line representing the regression equation

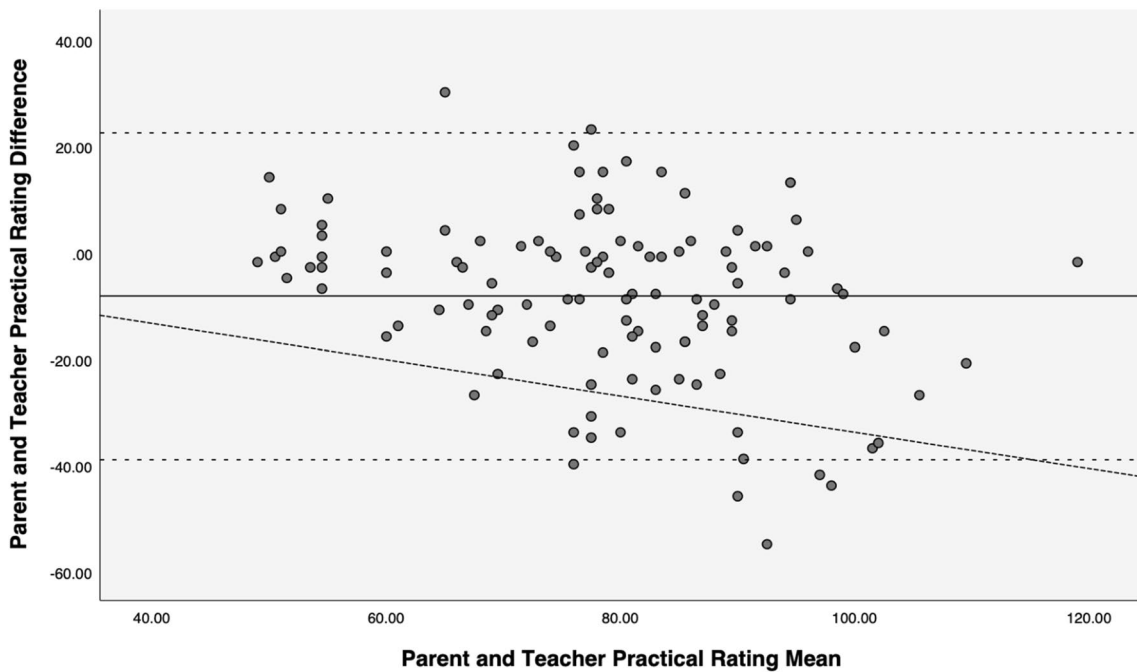


Fig. 4 Bland–Altman plot of parent/caregiver-teacher scores on the ABAS-3 practical domain. Scatter plot with a light gray background and dark gray data points representing parent-teacher practical difference scores regressed onto the mean scores. The plot includes one

solid black line representing the mean difference score, two dotted lines representing the limits of agreement (or confidence intervals), and one diagonal, dotted line representing the regression equation

Table 3 Regression analyses: variable predicting parent-teacher GAC, conceptual, social, and practical difference scores

Variable	GAC			Conceptual			Social			Practical		
	B	SE	<i>p</i>	B	SE	<i>p</i>	B	SE	<i>p</i>	B	SE	<i>p</i>
Constant	14.95	10.38	0.15	20.68	7.84	0.01*	5.96	5.33	0.27	16.06	11.20	0.16
Age	−0.06	0.04	0.15	−0.07	0.04	0.08	−0.06	0.04	0.14	−0.06	0.04	0.18
Diagnosis	1.55	1.90	0.42	–	–	–	–	–	–	3.15	2.05	0.13
Verbal IQ	−0.08	0.10	0.41	−0.12	0.07	0.10	–	–	–	−0.08	0.11	0.47
SCQ-lifetime	−0.70	0.24	0.01*	−0.54	0.23	0.02*	−0.51	0.24	0.04*	−0.86	0.26	<0.01*

**p* < 0.05

paired GAC scores fell within the LOA, which suggested acceptable agreement across all pairs. There were a few outliers on the Conceptual, Social, and Practical plots, but many of the paired scores fell within the upper and lower LOA. In the Bland–Altman regression analyses, the mean scores were regressed onto the difference scores. There were significant relationships between the difference and means for the GAC composite ($B = -0.377$, $t = -3.846$, $p < 0.001$), Conceptual ($B = -0.331$, $t = -3.360$, $p < 0.001$), Social ($B = -0.479$, $t = -4.131$, $p < 0.001$), and Practical domain ($B = -0.341$, $t = -3.455$, $p = 0.001$).

Predictors of Parent Teacher Differences

In addition to the analysis of the Bland–Altman plots, potential predictors of the difference scores across the total sample for the GAC, Conceptual, Social, and Practical domains using a linear regression model. The variables included were participant age at the time of their evaluation, verbal and nonverbal IQ, autism severity (represented by parent report on the SCQ-Lifetime), parent education, and final diagnosis or diagnoses (Table 3). There were several significant correlations between the parent-teacher difference scores and the included variables (refer to Online Resource 5 for all correlations). Nonsignificant correlations were found between parent-teacher difference and nonverbal IQ and parent education; thus, those two variables were not included in the final regression model. The parent-teacher difference scores on the GAC were predicted by the SCQ-Lifetime scores, which explained 30% of the variance in GAC difference scores ($B = -0.70$, $p < 0.05$). Specifically, as SCQ-Lifetime scores increased, parent-teacher difference scores decreased by 0.70 on the GAC. SCQ-Lifetime scores also predicted differences on informants' ratings on the Conceptual ($B = -0.54$, $p < 0.05$), Social ($B = -0.51$, $p < 0.05$), and Practical domains ($B = -0.86$, $p < 0.01$). Conceptual, Social, and Practical scores explained 46, 49, and 13% of the variance in parent-teacher difference scores respectively.

Discussion

Given the increasing prevalence rates for ASD and other neurodevelopmental disorders, clinicians require access to measures that allow them to identify client's current functioning as well as measures that provide useful information for intervention planning. Adaptive behavior measures are necessary in evaluations, not only because they are needed for diagnostic purposes, but because of the need to identify skills for intervention. Gathering data from multiple informants provides information about the client's performance across settings (e.g., school, home), which is consistent with best practice in evaluation (Floyd et al., 2015). While discrepancies on behavior measures are common for youth with ASD (e.g., Stratis et al., 2015), they are potentially problematic as they complicate clinical decision-making including clinical interpretation, differential diagnosis, treatment planning, and evaluating treatment efficacy (Hawley & Weisz, 2003). Research has identified discrepancies between parent and teacher informants when rating behavior in youth with ASD (Dickson, 2018; Jordan et al., 2019; McDonald et al., 2016), and less often, with children with ID (Ellison, 2016). It is important for clinicians and other stakeholders to understand what patterns of discrepancies might be encountered in clinical settings, and to recognize what individual factors might be more likely to result in such discrepancies. Thus, we aimed to determine the extent to which parents and teachers agreed on a measure of adaptive behavior for youth with ASD, ID, and/or GDD. Specifically, we asked the following research questions: (1) What is the level agreement between parent and teacher ratings on the ABAS-3? (2) Are discrepancies between informants associated with the child's diagnosis? (3) What are the trends in the magnitude of discrepancies across the entire sample? and (4) Do client factors predict parent/teacher discrepancies on the ABAS-3? The current study adds to this line of research by including a sample of participants with diverse cognitive abilities across a wide age range and aimed to better understand factors that might predict rating discrepancies across multi-informants.

Parent-Teacher Agreement

With regard to the level of agreement between parent and teacher ratings (i.e., research question one), while the paired ratings were significantly and positively correlated, there was poor reliability. This was observed for the GAC and all three domains on the ABAS-3. This finding is similar to previous research by Jordan et al. (2019) where positive associations between parent-teacher informants were identified on the GAC and Practical domain. However, in the current study, the magnitude of the relationship between parent-teacher GAC, Conceptual, and Social scores was slightly larger than those in Jordan et al. (2019). Unlike previous research, reliability was interpreted in the current study using Koo and Li's (2016) guidelines for intraclass correlation coefficients (ICC) where 0.70 is considered acceptable reliability between two measures. Relying solely on correlations between informants may be insufficient for measuring agreement, because correlational data only provides insight to the magnitude of the relationship between variables. The poor reliability found between parent-teacher informants in the current study suggested that parent and teacher informants were not reliable with each other in reporting youth adaptive behavior, which warranted further investigation.

Parent-Teacher Discrepancies Across Subsamples

Discrepancies between informants were further examined and compared to investigate if there was an association between diagnoses and patterns of discrepancies (i.e., research question two). Teachers rated all youth in our sample as having *more* adaptive behavior skills compared to parent report on all ABAS-3 composite and subdomain scores. These results are different from previous research where parent-teacher discrepancies were only found for the GAC and Conceptual rating (Jordan et al., 2019). In the current study, subsamples were examined to determine how parents and teachers differ in rating adaptive behavior for youth with ASD only and youth with co-occurring ID or GDD. Teachers rated participants with ASD ($n=91$) with more General Adaptive, Conceptual, Social, and Practical skills than their parents. There were no significant differences between informants rating the subsample of participants with ID and co-occurring diagnoses, except on the Social domain. However, all difference scores on the ABAS-3 were negative, which indicated higher teacher ratings and scores. This may be an artifact of the small sample size ($n=24$), but this finding is consistent with previous research demonstrating more agreement between informants' ratings with decreasing IQ (Stratis & Lecavalier, 2015) and suggest that parents and teachers are more likely to agree on skill development in autistic youth with ID as they demonstrate clearer deficits in adaptive behavior. Parents and teachers rating participants

with ASD and ID/GDD disagreed on level of social skills. The assessment of social skills may be more difficult due to amount of subjective judgement used (Lecavalier & Butter, 2010). However, this finding is somewhat surprising given research that suggests impairments in social skills may be more easily observed in youth with ASD and ID (Wilkins & Matson, 2009). Nonetheless, teachers consistently rated youth with ASD, ID, and/or GDD with more adaptive behavior skills. This may be related to the differences across settings (i.e., home versus school). Specifically, teachers may have more opportunities to observe social skills as they can readily compare individual students to their same-aged peers. Teachers are highly effective when identifying friendships, peer rejection, and social concerns in their classroom (e.g., Lane & Menzies, 2005), which is likely explained by their exposure and direct interaction with multiple children in one school year as well as across their teaching career.

Trends Across the Total Sample

Following the data analysis plan used in previous research (Jordan et al., 2019; McDonald et al., 2016), the current study evaluated the magnitude of parent-teacher discrepancies across the total sample's range of scores (i.e., research question three). There were significant, negative relationships between the parent-teacher difference scores and mean difference score for the GAC and three ABAS-3 domains. Thus, differences in parent-teacher ratings differed significantly across the sample's range of scores, specifically, larger parent-teacher score differences were observed with higher teacher scores. All Bland–Altman plots had negative means (i.e., higher teacher scores) and downward trends (i.e., increasing negative difference scores). This is unlike previous research where there were no significant differences were found (Jordan et al., 2019), but this finding is somewhat similar to McDonald et al (2016) who found positive, significant relationships between informant difference scores and mean scores. However, results from the current study are different than McDonald et al. (2016) in that—in the current study—teachers rated youth higher on the ABAS-3 whereas parents in previous research rated autistic youth without ID as having more adaptive skills on the BASC-2. This may represent the differences in sample characteristics and measure selection. This discrepancy across individual informant pairs (i.e., *t*-test) as well as all pairs included in the current study (i.e., Bland–Altman) may represent differences in informants' understanding of adaptive behavior as well as adaptive behavior expectations.

Predicting Parent-Teacher Discrepancies

The current study extended on previous research investigating child and family factors that explain discrepancies

between multiple informants on behavior rating scales of youth with ASD, ID, and GDD (Dickson, et al., 2018; Jordan et al., 2019; McDonald et al., 2016; Stratis & Lecavalier, 2015). Several child factors were significantly associated with parent-teacher difference scores (i.e., research question four). The strength of the relationships between age, diagnosis, verbal IQ, and family income were small ($r=0.21-0.28$), so it is unclear if this finding has clinical significance. However, autism severity scores on the SCQ-Lifetime had moderate, significant relationships with parent-teacher rating discrepancies for every ABAS-3 domain with correlations ranging from 0.27 to 0.39. This finding is in contrast to studies using more comprehensive measures of ASD. For example, Jordan et al. (2019) did not find any significant relationships in a study including participants' scores on the ABAS-3 and *Autism Diagnostic Interview-Revised* (ADI-R; Rutter et al., 2003a, 2003b), which was used to characterize autism symptomology. Similarly, comparison scores from the ADOS-2 were not related to parent-teacher discrepancies on the Vineland-II (Dickson et al., 2018; Kanne et al., 2011). Participant ratings on the SCQ-Lifetime may be more related to informant discrepancies than other autism-specific measures because it is a screening instrument and is often completed by an informant who also completes other behavior measures including the ABAS-3.

Given the significant relationships found between ABAS-3 difference scores and selected child variables, regression analyses were used to determine predictors of parent-teacher discrepancies. SCQ-Lifetime was the only significant predictor of all ABAS difference scores, and specifically, as SCQ-Lifetime scores increased, the GAC, Conceptual, Social, and Practical parent-teacher difference scores decreased. A high score on the SCQ-Lifetime indicates an increased risk of having ASD and having higher scores on an autism screener predicted increased informant agreement in the current study. Informants may find it easier, and thus are more consistent, when rating youth with more obvious or apparent skill deficits. This pattern of clinical symptomology impacting informant agreement has been found in other clinical samples where multiple informants were likely to endorse symptoms when clients have demonstrated elevated ASD symptom severity (e.g., Carlson & Youngstorm, 2003).

Clinical Implications

Discrepancies between informant ratings may be the result of several factors including the informants, the settings in which the informant and client interact, as well as the client themselves. First, it is well understood and agreed that parents and teachers have different relationships and roles in a child's life. Parents and teachers interact with children in different settings, which reflects contextual variations that

naturally disrupts agreement on behavior measures. Therefore, parents and teachers may have varied opportunities to observe adaptive behaviors. Teachers can observe and compare one student to several, same-aged peers, which may enhance their ability to identify strengths and weaknesses in individual children. Research suggests that executive functioning skills predict differences in adaptive behavior (Tomaszewski et al., 2020), and, in some cases, more than IQ (Puliese et al., 2015). These related skills and their impact on adaptive behavior may be more readily observed in schools due to the demands in that setting. However, parents might be better positioned to speak to a client's current daily living skills (e.g., feeding, hygiene) as well as their development of these skills overtime. One potential explanation for the higher teacher ratings of adaptive skills found in the current study could be that the structures, routines, and supports of school settings contribute to the performance of more adaptive behavior skills at school than at home (which may be less structured environments). Moreover, aspects of the settings like time of day (e.g., parents are more likely to see their child when they are tired) and other conditions may result in discrepant ratings.

The ABAS-3 manual offers other possible explanations behind informant discrepancies including misunderstanding of the instrument, and errors when completing the rating form. Differences in informant experience and training may further influence ratings. Notably, one informant may have enhanced observations skills or preparation for completing the rating form. Relatedly, it is also likely that parents and teachers have a different understanding of adaptive behavior. Informant characteristics like reading skill, language use, and cultural differences may impact understanding and expectations for adaptive behavior as well as accurate completion of a measure.

Individual client characteristics also contribute to discrepant ratings which has implications for differential diagnosis. In the current study, increased autism symptom severity (as measured by the SCQ-lifetime) predicted informant agreement on the ABAS-3. In addition to autism symptomology, Kanne et al. (2011) found significant relationships between adaptive behavior and age, and IQ in youth with autism. Specifically, youth with ASD showed more deficits in adaptive behavior when they were older, and youth with average cognitive abilities demonstrated larger differences in adaptive behavior compared to youth with ID. Clients with lower cognitive abilities are more likely to show commensurate adaptive behavior, whereas clients with higher IQ demonstrate discrepancies between their cognitive and adaptive behavior abilities. Moreover, the current study suggests that parent-teacher data obtained through adaptive behavior measures is more discrepant when youth show *less* autism symptomology. Clients with subtle autism presentations and solid cognitive abilities may experience adaptive behavior

deficits, but in these instances informants may strongly disagree on adaptive behavior rating scales further complicating identification of client strengths, weaknesses, and treatment planning.

Researchers have recommended that clinicians use what they know about informant discrepancies to anticipate informant differences and then to use this information while constructing their assessment plan (De Los Reyes et al., 2015). Clinicians can maximize their time spent conducting family interviews and/or use other assessments to gather more information on adaptive behavior to better understand these differences and the child's overall adaptive behavior functioning. The ABAS-3 manual also provides guidance to clinicians on how to navigate such discrepancies by calculating the difference score and its statistical significance when informant differences are present (Harrison & Oakland, 2015). However, this solution may not be feasible due to time constraints during clinical evaluations. Thus, seeking out additional information with other assessment tools may be more useful for diagnoses and intervention planning. If the clinician assumes that results on the ABAS-3 will look different between parents and teachers, the clinician could plan to include questions related to settings differences (e.g., opportunities to engage in daily living skills, behavior management practices, consequences). Moreover, direct assessment of adaptive behavior is recommended as a necessary supplement to measures and interviews (Floyd et al., 2015). Behavior observations are advantageous as they allow the clinician to observe clients' behaviors in structured and unstructured settings and evaluate their response to different demands.

Limitations and Future Directions

Although the current study has potential implications for research and practice, results should be considered within the context of their limitations. First, the current study found differences in parent and teacher reports for the small subsample of youth with ID, GDD, and co-morbid ASD ($n = 24$) only for the Social domain on the ABAS-3. However, there may be differences in other domains across informant reports that were not detected as a result of the small subsample size and low power. Second, the sample was limited to one geographical area in the United States (i.e., Midwest) and the sample largely consisted of White (83%), male (83%) clients diagnosed with ASD only (79%). Girls and youth of color are more likely to be undiagnosed with ASD compared to their male, White peers (Wiggins et al., 2020), which is evident in the current study as well. Disparities in ASD diagnoses have been attributed to many factors, such as levels of stigma and providers dismissing family's concerns (Stahmer et al., 2019). When it comes to measurement of children's behavior, previous research indicates that informant

agreement may vary based on the race and ethnicity of the child. For example, Black parents were less likely to report concerns regarding their child's ASD symptoms on screening measures (Donohue et al., 2017). However, in a separate study, Black children were more likely than White children to fall in the at-risk range on an autism screening tool again, after their parents participated in the follow-up interview that included more questions about autism symptomology (Dai et al., 2021).

Additionally, some research suggests that youth of color with ASD (specifically Black and Hispanic/Latinx youth) demonstrated more severe behaviors and language delays when compared to their White counterparts (Angell et al., 2018). These findings suggest increased severity of ASD symptomology in Black and Latinx youth. As noted earlier, informants tend to be more consistent in their ratings when youth demonstrate more obvious deficits, and the current study showed reduced ASD symptomology predicted greater disagreement across reporters. Given this, informants may demonstrate more agreement on behavior measures when rating youth of color who demonstrate increased severity in symptoms. As such, if the sample in the current study was more racially and ethnically diverse, the reports between parent and teacher informants may have been less discrepant.

In the area of informant discrepancies on adaptive behavior measures, research suggests differences across several individual factors such as age (Hill et al., 2015), diagnoses (Matson et al., 2009), and IQ (Liss et al., 2001). For example, youth with intellectual disabilities have shown more adaptive behavior skills than youth with ASD (Kanne et al., 2011), which also has been found in samples of youth living outside of the United States (e.g., Alvares et al., 2020; Kilincaslan et al., 2019). However, research examining differences in informant ratings of adaptive behavior across youth from diverse racial and ethnic backgrounds is lacking. In a study with a larger, diverse sample, parents from different racial and ethnic background report autism symptomology and concern differently (Azad et al., 2021), so it is likely that there are differences on adaptive behavior measures too. Future research should replicate the current study using a more diverse sample, and in addition to the comprehensive data analysis plan used in the current study, other methods of understanding individual, familial, and cultural factors is encouraged (i.e., culturally relevant assessment practices).

Other limitations surround issues of assessment, measurement, and methodology. More specifically, the current study did not control for type of test used to measure participants' verbal and nonverbal IQ. In addition, the current study is cross-sectional, and data included in the current study were gathered retrospectively. Studies using alternative methodology could provide further insight to the nature of the relationships among patient characteristics and informant agreement of adaptive behavior skills over time. Fourth, the

current study only examined informant agreement on one measure of adaptive behavior and the degree to which these findings generalize to other measures is unknown. Future researchers might compare these findings to informant agreement on other scales (e.g., *Adaptive Behavior Evaluation Scale, Third Edition*; McCarney & House, 2017) and possibly evaluate the clinical utility of other sources of clients' adaptive behavior (e.g., individualized evaluation plan goals and progress, occupational therapy case notes). Finally, a review of the original research on the development of the ABAS-3 suggests that its psychometric rigor could have been more robust. The original psychometric research on the ABAS-3 showed low to moderate correlations between parent and teacher forms. As such, studies using the ABAS-3 (including the current study), are limited by the psychometric properties of the instrument.

Conclusion

In the current study, large discrepancies between parent-teacher informants were found on the ABAS-3 for youth with ASD, ID, and GDD. This may be a particularly important consideration for clinicians when conducting evaluations of patients with less severe autism symptomology and cognitive impairments. It is essential that clinicians have access to reliable and valid assessment tools that provide a snapshot of clients' present functioning as well as provide data that can be used to guide treatment and interventions; however, clinicians should expect there to be discrepancies between parents and teachers on adaptive behavior measures. Obtaining data from multiple informants is best practice and necessary to make diagnostic decisions, but based on the results of the present study, practitioners are encouraged to seek further information from parent-teacher informants (e.g., interviewing, behavior observations, etc.) when there are discrepancies to support clinical impressions and treatment planning.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10803-022-05550-x>.

Author Contributions All the listed authors made a significant contribution to the study. Stevens led and coordinated the current study's conceptualization, method, data collection and analysis, as well as ran the analyses, created the tables and figures, and contributed to each section within the manuscript. Selders supported the data collection and organization, contributed to several sections of the manuscript, and participated in the revision process. Jeckel contributed to data collection and organization, the introduction, and participated in the revision process. Brownfield contributed to data collection and organization. Nowell contributed to the manuscript by providing supervision throughout the study timeline (in the initial conceptualization of the study, data analysis plan, etc.) and reviewed the final draft of the manuscript.

References

- Altman, D. G., & Bland, J. M. (1983). Measurement in medicine: The analysis of method comparison studies. *Journal of the Royal Statistical Society*, 32(3), 307–317.
- Alvares, G. A., Bebbington, K., Cleary, D., Evans, K., Glasson, E. J., Mayberry, M. T., Pillar, S., Uljarvić, M., Varcin, K., Wray, J., & Whitehouse, A. J. O. (2020). The misnomer of 'high functioning autism': Intelligence is an imprecise predictor of functional abilities at diagnosis. *Autism*, 24(1), 221–232. <https://doi.org/10.1177/1362361319852831>
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). American Psychiatric Association.
- 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. <https://doi.org/10.1016/bs.irrdd.2018.08.003>
- Azad, G., Hologue, C., Pfeiffer, D., Dillon, E., Reetzke, R., Kalb, L., Menon, D., Hong, J. S., & Landa, R. (2021). The influence of race on parental beliefs and concerns during an autism diagnosis: A mixed-method analysis. *Autism*. <https://doi.org/10.1177/13623613211044345>
- Baio, J., Wiggins, K., Christensen, D. L., Maenner, M. J., Daniels, J., Warren, Z., Kurzius-Spencer, M., Zahorodny, W., Rosenberg, C. R., White, T., Dunkin, M. S., Imm, P., Nikolaou, L., Yeargin-Allsopp, M., Lee, L., Harrington, R., Lopez, M., Fitzgerald, R. T., Hewitt, A., ... Dowling, N. P. (2018). Prevalence of autism spectrum disorder among children ages 8 years—autism and developmental disabilities monitoring network, 11 sites, United States. *Surveillance and Summaries*, 67(6), 1–23. <https://doi.org/10.15585/mmwr.ss9706a1>
- Bland, J. M., & Altman, D. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *The Lancet*, 327(8476), 307–310. [https://doi.org/10.1016/S0140-6736\(86\)90837-8](https://doi.org/10.1016/S0140-6736(86)90837-8)
- Carlson, G. A., & Youngstrom, E. A. (2003). Clinical implications of pervasive manic symptoms in children. *Society of Biological Psychiatry*, 53, 1150–1158. [https://doi.org/10.1016/S0006-3223\(03\)00068-4](https://doi.org/10.1016/S0006-3223(03)00068-4)
- Cohen, J. E. (1988). *Statistical power analysis for the behavioral sciences*. Lawrence Erlbaum Associates Inc.
- 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>
- De Los Reyes, A. (2011). Introduction to the special section: More than measurement error: Discovering meaning behind informant discrepancies in clinical assessments of children and adolescents. *Journal of Clinical Child Adolescent Psychology*, 40(1), 1–9. <https://doi.org/10.1080/15374416.2011.533405>
- De Los Reyes, A., Augenstein, T. M., Wang, M., Drabick, D. A. G., Burgers, D. E., & Rabinowitz, J. (2015). The validity of multi-informant approach to assessing child and adolescent mental health. *Psychological Bulletin*, 141(4), 858–900. <https://doi.org/10.1037/a003848498>
- Dickson, K. S., Suhrheinrich, J., Rieth, S. R., & Stahmer, A. C. (2018). Parent and teacher concordance of child outcomes for youth with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, 48(5), 1423–1435. <https://doi.org/10.1007/s10803-017-3382-z>

- Donohue, M. R., Childs, A. W., Richards, M., & Robins, D. L. (2017). Race influences parent report of concerns about symptoms of autism spectrum disorder. *Autism*, 23(1), 100–111. <https://doi.org/10.1177/1362361317722030>
- Elliot, C. D. (2007). *Differential ability scales-II (DAS-II)*. Pearson.
- Ellison, K. S., Bundy, M. B., Wygant, D. B., & Gore, J. S. (2016). Don't forget about the teachers! Parent and teacher BASC-2 reports of children on the DSM-5 autism spectrum. *Journal of Mental Health Research in Intellectual Disabilities*, 9(3), 119–132. <https://doi.org/10.1080/19315864.2016.1181812>
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39, 175–191.
- Floyd, R. G., Shands, E. I., Alfonso, V. C., Phillips, J. F., Autry, B. K., Mosteller, J. A., Skinner, M., & Irby, S. (2015). A systematic review and psychometric evaluation of adaptive behavior scales and recommendations for practice. *Journal of Applied School Psychology*, 31(1), 83–113. <https://doi.org/10.1080/15377903.2014.979384>
- Harrison, P. L., & Oakland, T. (2015). *Adaptive behavior assessment system, third edition (ABAS-3)*. Western Psychological Services.
- Hattier, M. A., Matson, J. L., Belva, B. C., & Adams, H. (2013). The effects of informant age and education level on childhood ASD symptom endorsement. *Research in Autism Spectrum Disorders*, 7(2), 398–404. <https://doi.org/10.1016/j.rasd.2012.11.001>
- Hawley, K. M., & Weisz, J. R. (2003). Child, parent and therapist (dis) agreement on target problems in outpatient therapy: The therapist's dilemma and its implications. *Journal of Consulting and Clinical Psychology*, 71(1), 62–70. <https://doi.org/10.1037/0022-006X.71.1.62>
- Hill, T. L., Gray, S. A. O., Kamps, J. L., & Varela, R. E. (2015). Age and adaptive functioning in children and adolescents with ASD: The effects of intellectual functioning and ASD symptom severity. *Journal of Autism and Developmental Disorders*, 45, 4074–4083. <https://doi.org/10.1007/s10803-015-2522-6>
- Hyman, S. L., Levy, S. E., Myers, S. M., Council on Children with Disabilities, Section on Developmental and Behavioral Pediatrics. (2020). Identification, evaluation, and management of children with autism spectrum disorder. *Pediatrics*, 145(1), 1–64. <https://doi.org/10.1542/peds.2019-3447>
- Jordan, A. K., Thomeer, M. L., Lopata, C., Donnelly, J. P., & McDonald, C. A. (2019). Informant discrepancies in the assessment of adaptive behavior of children with autism spectrum disorder. *Journal of Autism & Developmental Disorders*, 5, 2024–2034. <https://doi.org/10.1007/s10803-018-03876-z>
- Kanne, S. M., Gerber, A. J., Quirnbach, L. M., Sparrow, S. S., Cicchetti, D. V., & Saulnier, C. A. (2011). The role of adaptive behavior in autism spectrum disorders: Implications for functional outcome. *Journal of Autism and Developmental Disorders*, 41(8), 1007–1018. <https://doi.org/10.1007/s10803-010-1126-4>
- Kilincaslan, A., Kocas, S., Bozkurt, S., Kaya, I., Derin, S., & Aydin, R. (2019). Daily living skills in children with autism spectrum disorder. And intellectual disability: A comparative study from Turkey. *Research in Developmental Disabilities*, 85, 187–196. <https://doi.org/10.1016/j.ridd.2018.12.005>
- Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, 15(2), 155–163. <https://doi.org/10.1016/j.jcm.2016.02.012>
- Lane, K. L., & Menzies, H. M. (2005). Teacher-identified students with and without academic and behavioral concerns: Characteristics and responsiveness. *Behavioral Disorders*, 31(1), 65–83. <https://doi.org/10.1177/019874290503100103>
- Lecavalier, L., & Butter, E. M. (2010). Assessment of social skills and intellectual disability. *Practitioner's guide to empirically based measures of social skills* (pp. 172–192). Springer.
- Liss, M., Harel, B., Fein, D., Allen, D., Dunn, M., Feinstein, C., Morris, R., Waterhouse, L., & Rapin, I. (2001). Predictors and correlates of adaptive functioning in children with developmental disorders. *Journal of Autism and Developmental Disorders*, 31, 219–230. <https://doi.org/10.1023/A:1010707417274>
- Lord, C., Rutter, M., DiLavore, P. C., Gotham, K., & Bishop, S. (2012). *Autism diagnostic observation schedule, second edition (ADOS-2)*. Western Psychological Services.
- 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., Hutson, R., ... Dietz, D. M. (2020). Prevalence of autism spectrum disorder among children aged 8 years. Autism and developmental disabilities monitoring network, 11 sites, United States, 2016. *Surveillance and Summaries*, 69(4), 1–12. <https://doi.org/10.15585/mmwr.ss6904a1>
- McCarney, S. B., & House, S. N. (2017). *Adaptive behavior evaluation scale, third edition (ABES-3)*. Hawthorne Educational Services Inc.
- McDonald, C. A., Lopata, C., Donnelly, J. P., Thomeer, M. L., Rodgers, J. D., & Jordan, A. K. (2016). Informant discrepancies in externalizing and internalizing symptoms and adaptive skills of high-functioning children with autism spectrum disorder. *School Psychology Quarterly*, 31(4), 467–477. <https://doi.org/10.1037/spq0000150>
- Mullen, E. M. (2001). *Mullen scales for early learning (Mullen)*. Pearson.
- Pugliese, C. E., Anthony, L., Strang, J. F., Dudley, K., Wallace, G. L., & Kenworthy, L. (2015). Increasing adaptive behavior skill deficits from childhood to adolescence in autism spectrum disorder: Role of executive function. *Journal of Autism and Developmental Disabilities*, 45, 1579–1587. <https://doi.org/10.1007/s10803-014-2309-1>
- Ranganathan, P., Pramesh, C. S., & Aggarwal, R. (2017). Common pitfalls in statistical analysis: Measures of agreement. *Perspectives in Clinical Research*, 8(4), 187–191. https://doi.org/10.4103/picr.PICR_123_17
- Reynolds, C. R., & Kamphaus, R. W. (2004). *Behavior assessment system for children, second edition (BASC-2)*. Pearson.
- Roid, G. H., Miller, L. J., Pomplin, M., & Kock, C. (2013). *Leiter international performance scale, third edition (Leiter-3)*. Western Psychological Services.
- Rutter, M., Bailey, A., & Lord, C. (2003a). *SCQ. The social communication questionnaire*. Western Psychological Services.
- Rutter, M., LeCouteur, A., & Lord, C. (2003b). *Autism diagnostic interview—revised (ADI-R)*. Western Psychological Services.
- Saulnier, C. A., & Klaiman, C. (2018). *Essentials of adaptive behavior assessment of neurodevelopmental disorders*. Wiley.
- 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., Shogren, K. A., & Yeager, M. H. (2010). *Intellectual disability: Diagnosis, classification, and systems of supports* (11th ed.). American Association on Intellectual and Developmental Disabilities.
- Sparrow, S. S., Cicchetti, D. V., & Balla, D. A. (2006). *Vineland adaptive behavioral scales, second edition (Vineland-II)*. Pearson.
- Stahmer, A. C., Vojnoska, S., Iadarola, S., Straiton, D., Segovia, F. R., Luelmo, P., Morgan, E. H., Lee, H. S., Javed, A., Bronstein, B., Hochheimer, S., Cho, E., Aranbarri, A., Mandell, D., Massrick, E. M., Smith, T., & Kasari, C. (2019). Caregiver voices: Cross-cultural input on improving access to autism services. *Journal of*

- Racial and Ethnic Health Disparities*, 6, 752–773. <https://doi.org/10.1007/s40615-019-00575-y>
- Stratis, E. A., & Lecavalier, L. (2015). Informant agreement for youth with autism spectrum disorder or intellectual disability: A meta-analysis. *Journal of Autism and Developmental Disorders*, 45(4), 1026–1041. <https://doi.org/10.1007/s10803-014-2258-8>
- Tassé, M. J., Schalock, R. L., Balboni, G., Bersani, H., Jr., Borthwick-Duffy, S. A., Spreat, S., Thissen, D., Widaman, K. F., & Zhang, D. (2012). The construct of adaptive behavior: Its conceptualization, measurement, and use in the field of intellectual disability. *American Journal of Intellectual and Developmental Disabilities*, 117(4), 291–303. <https://doi.org/10.1352/1944-7558-117.4.291>
- Tomaszewski, B., Hepburn, S., Blakeley-Smith, A., & Rogers, S. (2020). Developmental trajectories of adaptive behavior from toddlerhood to middle childhood in autism spectrum disorder. *American Journal of Intellectual and Developmental Disabilities*, 125(3), 155–169. <https://doi.org/10.1352/1944-7558-125.3.155>
- Wechsler, D. (2008). *Wechsler adult intelligence scale, fourth edition (WAIS-IV)*. Pearson.
- 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 (WPPSI-IV)*. Pearson.
- Wechsler, D. (2014). *Wechsler intelligence scale for children, fifth edition (WISC-V)*. Pearson.
- Wiggins, L. D., Durkin, M., Esler, A., Lee, L., Zahorodny, W., Rice, C., Yeargin-Allsopp, M., Dowling, N. F., Hall-Lande, J., Morrier, M. J., Christensen, D., Shenouda, J., & Baio, J. (2020). Disparities in documented diagnoses of autism spectrum disorder based on demographic, individual, and service factors. *Autism Research*, 13(3), 464–473. <https://doi.org/10.1002/aur.2255>
- Wilkins, J., & Matson, J. L. (2009). A comparison of social skills profiles in intellectually disabled adults with and without ASD. *Behavior Modification*, 33(2), 143–155. <https://doi.org/10.1177/0145445508321880>
- Zablotsky, B., Black, L. I., Maenner, M. J., Schieve, L. A., & Blumberg, S. J. (2015). Estimated prevalence of autism and other developmental disabilities following questionnaire changes in the 2014 national health statistics reports. *National Health Statistics Reports*, 87, 1–20.

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