RESEARCH ARTICLE

DAS-II Cognitive Profiles Are Not Diagnostically Meaningful For Autism: A ROC Analysis

Caitlin C. Clements, Timea Sparding, Robert T. Schultz, Benjamin E. Yerys, and Marley W. Watkins

Intelligence assessment is an integral part of a comprehensive autism evaluation. Many past studies have described a cognitive profile of autistic individuals characterized by higher nonverbal than verbal IQ scores. The diagnostic utility of this profile, however, remains unknown. We leveraged receiver operating characteristic methods to determine the sensitivity, specificity, and area under the curve (AUC) of three different IQ profiles in a large sample of children who have an autism spectrum disorder diagnosis (N = 1,228, Simons Simplex Collection) who completed the Differential Ability Scales—Second Edition (DAS-II), School Age compared to the normative sample provided by the DAS-II publisher (N = 2,200). The frequently discussed nonverbal > verbal IQ profile performed near chance at distinguishing ASD from normative individuals (AUC: 0.54, 95% CI [0.52–0.56]), and performed significantly worse for females than males (AUC: females: 0.46 [0.41–0.52]; males: 0.55 [0.53–0.58]). All cognitive profiles showed AUC < 0.56. We conclude that while significant differences between verbal and nonverbal IQ scores exist at the group level, these differences are small in an absolute sense and not meaningful at an individual level. We do not recommend using cognitive profiles to aid in autism diagnostic decision-making. ***Autism Res*** *2020, 13: 2143–2154.* © 2020 International Society for Autism Research, Wiley Periodicals LLC.

Lay Summary: Some researchers and clinicians have reported an "autistic cognitive profile" of higher nonverbal intelligence than verbal intelligence. In an analysis of over 1,000 autistic children, we found that the *group's* average nonverbal intelligence is usually higher than their verbal intelligence. However, this pattern should not be used by clinicians to make an *individual* diagnosis of autism because our results show it is not helpful nor accurate.

Keywords: autism; DAS-II; intelligence; cognitive profiles; ROC

Introduction

Autism spectrum disorder (ASD) often co-occurs with intellectual disability and other developmental disorders [Charman et al., 2011; Klinger & Dudley, 2019; Lerner, Mazefsky, White, & McPartland, 2018], which makes the differential diagnosis of ASD exceptionally challenging [Hayes, Ford, Rafeeque, & Russell, 2018; Huerta & Lord, 2012; Lord, Corsello, & Grzadzinski, 2014; McDonnell et al., 2019]. Accordingly, cognitive assessment has often been recommended as part of a comprehensive multidisciplinary assessment for ASD referrals [Aiello, Ruble, & Esler, 2017; Caterino, 2014; Filipek et al., 2000; Kroncke, Willard, & Huckabee, 2016; Ozonoff, Goodlin-Jones, & Solomon, 2005; Saulnier & Ventola, 2012; Volkman et al., 2014]. Some of the most commonly used cognitive assessments for ASD evaluations of school-age children include the DAS-II (Differential Ability Scales, 2nd Edition, Elliott, 2007a) and the current versions of the Wechsler scales such as the WISC-V [Wechsler Intelligence Scale for Children-5th Edition, Wechsler, 2014] and WASI-II [Wechsler Abbreviated Scales of Intelligence, 2nd Edition, Wechsler, 2011].

Cognitive Profiles in ASD

Studies of IQ scores among children with autism often found uneven cognitive profiles with nonverbal scores higher than verbal scores, regardless of the particular cognitive instrument used. A review of 16 published studies involving 333 ASD children tested with a Wechsler scale found an N-V difference of 9.3 points, that is, greater than 0.60 standard deviation difference [Lincoln, Courchesne, Allen, Hanson, & Ene, 1998]. Other individual studies frequently reported N > V with Wechsler scales [Giofrè et al., 2019; Jones & Lord, 2013; Kuriakose, 2014; Li, Du,

From the Center for Autism Research, The Children's Hospital of Philadelphia, Philadelphia, Pennsylvania (C.C.C., R.T.S., B.E.Y.); Psychology Department, University of Pennsylvania, Philadelphia, Pennsylvania (C.C.C.); Department of Psychiatry and Neurochemistry, Institute of Neuroscience and Physiology, The Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden (T.S.); Department of Pediatrics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania (R.T.S.); Department of Psychiatry, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania (B.E.Y.); Department of Educational Psychology, Baylor University, Waco, Texas (M.W.W.)

Received November 25, 2019; accepted for publication May 24, 2020

Address for correspondence and reprints: Caitlin C. Clements, The Children's Hospital of Philadelphia, Center for Autism Research, 2716 South St., Philadelphia, PA 19104. E-mail: clements@sas.upenn.edu

Published online 21 July 2020 in Wiley Online Library (wileyonlinelibrary.com)
DOI: 10.1002/aur.2336

Luan, Li, & Ousley, 2017; Mayes & Calhoun, 2008]. N > V score profiles have also been found among ASD children when other cognitive instruments were employed [DAS: Frazier, Georgiades, Bishop, & Hardan, 2014; Joseph, Tager-Flusberg, & Lord, 2002; Stanford Binet Intelligence Scales: Matthews et al., 2015; Leiter-R and KBIT-2: Scattone, Raggio, & May, 2012; Cognitive Assessment System: Taddei & Contena, 2013; for review of multiple instruments, see Tager-Flusberg & Joseph, 2004]. Studies employing a nonverbal cognitive instrument, the Leiter-Revised, also revealed a relative strength on visual processing tasks among children with ASD [Kuschner, Bennetto, & Yost, 2007; Mecca, Orsati, & de Macedo, 2014].

However, research results have not been consistent, and unique cognitive profiles have not always emerged from samples of children with ASD. Some studies reported no significant between-group differences in cognitive profile prevalence between children with and without ASD [Charman et al., 2011; Ehlers et al., 1997; Lennen, Lamb, Dunagan, & Hall, 2010; Siegel, Minshew, & Goldstein, 1996] and other studies found verbal scores superior to nonverbal scores [V > N; Ghaziuddin & Mountain-Kimchi,- 2004; Kanai et al., 2017]. Additionally, there has been little consideration of participants' sex although it may impact the reported prevalence of cognitive profiles [Ankenman, Elgin, Sullivan, Vincent, & Bernier, 2014; Duvall, Huang-Storms, Hill, Myers, & Fombonne, 2020; Howe et al., 2015; Nowell, Schanding, Kanne, & Goin-Kochel, 2015]. For example, in some samples, the N > V profile was less pronounced or absent among females [e.g., Ankenman et al., 2014].

Applying Research About Groups to Decisions About Individuals

Most critically, research on cognitive profiles among children with ASD has usually taken a discriminant validity approach [Elwood, 1993]. This approach is based on one of three types of group comparisons. In one method, the cognitive scores from a group of children with ASD are compared to the scores from a group of children without ASD. Statistically, significant mean N > V differences between groups are interpreted as evidence that the N > V cognitive profile is a valid sign of ASD [Lennen et al., 2010]. Please note that in many studies of ASD, the reported sample mean nonverbal IQ is greater than the sample mean verbal IQ [e.g., Nowell et al., 2015]. In another method, a cut score is determined a priori (e.g., 12–16 points) for significant nonverbal–verbal score discrepancies among a group of children with ASD; the proportion of children below/above the cut score in each group is statistically compared to infer discriminant validity [Tager-Flusberg & Joseph, 2004; Nowell et al., 2015]. In a third and final method, a significant mean difference between nonverbal and verbal scores within a group of children with ASD is taken as an indicator of ASD [Mayes & Calhoun, 2008]. Based on these group approaches, the N > V profile has been called a "useful autism-related phenotype" [Chapman et al., 2011, p. 65], a "valuable diagnostic tool for differential diagnosis" [Nader, Jelenic, & Soulières, 2015, p. 1], and a "potential sign that an autism diagnosis may be relevant" [Kroncke et al., 2016, p. 144].

However, results of nomothetic research cannot confidently be applied to idiographic decisions [Weiner, 2003]. Group differences speak to the typical member of a group, not to each individual within that group [Kraemer, Frank, & Kupfer, 2011]. Meehl and Rosen [1955] warned clinicians against using psychometric patterns or cut scores to aid diagnostic decision-making when the cut score is based on *between-group differences*, instead of on the cut score's accuracy in making *individual correct decisions*. More recently, this caution was specifically addressed to the use of cognitive profiles derived from groups of autistic children [Mandy, Murin, & Skuse, 2015]. Thus, when studying a nonverbal–verbal difference or any proposed cognitive profile cut score, the utility of the cut score should be evaluated in terms of decisions about individuals, not overall group differences.

Decisions About Individuals

As illustrated in Figure 1, there are four possible outcomes when cut scores are used to make decisions about individuals: true positive, false positive, true negative, and false negative. Thus, there are two types of correct decisions (true positive and true negative) as well as two types of errors (false positive and false negative). Several decision accuracy ratios can be computed from these four outcomes, including sensitivity (probability of ASD cognitive profile, given ASD diagnosis), specificity (probability of ordinary profile, given non-ASD diagnosis), positive predictive power (probability of ASD diagnosis, given ASD cognitive profile), and negative predictive power (probability of non-ASD diagnosis, given ordinary cognitive profile). Unfortunately, these ratios are influenced by the disorder prevalence as well as the cut score used in their computation [McFall & Treat, 1999; Swets, 1996; Youngstrom, 2014].

To ameliorate these limitations, a common metric for quantifying assessment information that is insensitive to both base rate and changing cut scores has emerged from signal detection theory [McFall & Treat, 1999; Swets, 1996; Youngstrom, 2014]. Originally developed to quantify how well a radar receiver detected electronic signals in the presence of noise, the receiver operating characteristic (ROC) is now widely used across many disciplines and is typically displayed as a graph with the sensitivity of each possible cut score plotted on the *y*-axis and the specificity of each possible cut score plotted on the *x*-axis [Youngstrom, 2014]. The resulting ROC graph

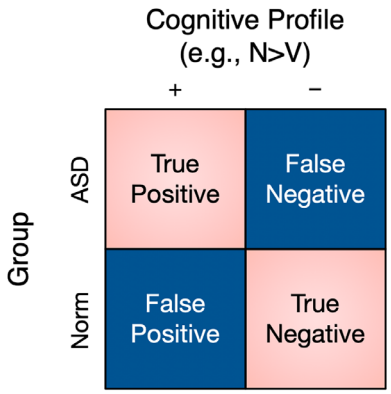

Figure 1. Decision matrix to classify an individual into the ASD or Norm group based on the presence (+) or absence (−) of a particular cognitive profile, such as nonverbal > verbal (N > V). Note. Sensitivity, specificity, positive predictive value, and negative predictive power are calculated using values in this matrix. ASD = children with autism spectrum disorder from the Simons Simplex Collection, Norm = the Differential Ability Scales—Second Edition School Age normative sample. Sensitivity = true positive ÷ true positive + false negative, specificity = true negative ÷ true negative + false positive, positive predictive power = true positive ÷ true positive + false positive, negative predictive power = true negative ÷ true negative + false negative.

visually illustrates how decision accuracy varies as the cut score is systematically moved across all possible values (see Figure 2).

The proportion of the graph's area that lies under the curve (AUC) quantifies overall decision accuracy. The AUC can be interpreted in terms of individual decisions: in this case, the probability that a randomly selected individual with ASD would exhibit a larger DAS-II N-V score discrepancy than a randomly selected individual without ASD. AUC values can range from 0.50 for decision accuracy at chance levels (displayed as a diagonal line in the ROC graph) to 1.00 for decisions that are perfectly accurate. Guidelines for decision accuracy posit that AUC values of 0.50 to 0.70 represent low accuracy, values of 0.70–0.90 indicate moderate accuracy, and values greater than 0.90 evince high accuracy [Swets, 1996]. However, AUC values in the 0.70s are probably more clinically informative given the reliability and validity of psychological instruments [Jarrett, Van Meter, Youngstrom, Hilton, & Ollendick, 2018; Youngstrom, 2014].

Only one study has employed ROC methods to ascertain the decision accuracy of cognitive profiles among children on the autism spectrum [Styck, Aman, & Watkins, 2019]. This study compared composite scores from a Wechsler child scale between the 2,200 children and adolescents in the test's normative sample and 79 school-aged students with ASD and found an AUC value of 0.62 for a cut score based on discrepant General Ability Index and Cognitive Proficiency Index (GAI > CPI). As with other studies, the ASD sample was small and unrepresentative in that it was collected from school records based on school evaluations, was limited to children enrolled in public education in the southwestern United States, and excluded children with comorbidities such as attention-deficit/hyperactivity disorder (ADHD). Additionally, Wechsler scales have been criticized for requiring relatively high levels of verbal comprehension for understanding test directions [Cormier, Wang, & Kennedy, 2016; Klinger, Mussey, & O'Kelley, 2018].

Current Study

The field of ASD diagnosis and cognitive profiles would benefit from an ROC analysis of cognitive assessment data from a large representative sample of autistic children. The DAS-II [Elliott, 2007a] has been recommended for cognitive assessments of children with ASD because of its ease of use, reliability, theoretical structure, reduced reliance on expressive language, and use of teaching items [Hausman-Kedem et al., 2018; Jones & Lord, 2013; Klinger et al., 2018; Kroncke et al., 2016; Saulnier & Ventola, 2012]. Thus, a ROC comparison of the DAS-II normative sample and a large national sample of autistic children might assist in clarifying the relative usefulness of cognitive profiles. These two large samples would also allow an investigation of the influence of child sex on cognitive profile results, and restriction of analyses to school-aged children would lessen the potential confounding influence of age. Compared to the one existing prior study, the current sample is nearly 20× larger, more geographically representative as participants were recruited from cities across the entire United States, and had more generous inclusion criteria for comorbidities (i.e., other psychological disorders such as ADHD were not exclusionary).

Method

Participants

The norm sample consisted of the nationally representative DAS-II School Age (ages 7–17 years) normative sample (n = 2,200) and was provided by Pearson, publisher of the DAS-II. The normative sample was equally divided between male and female participants with a mean age of 12.5 years (SD = 3.2 years). For additional information on this sample, see the DAS-II Technical Manual [Elliott, 2007b]. Briefly, recruitment was intended to capture a broad range of IQ levels via a nationally representative sample of children. Some recruitment efforts specifically targeted children with exceptionally high or exceptionally low IQ scores.

The ASD sample was drawn from the Simons Simplex Collection (SSC), which was a national multi-site study of 2,110 children aged 4–18 years who met gold-standard diagnostic criteria for ASD and had no immediate family

members with ASD (i.e., "simplex"). Participants completed a comprehensive diagnostic and behavioral battery that included the DAS-II. SSC participants were included in the present study if they were aged 7–17 years and were not missing any of the DAS-II School Years composite scores (n = 1,228, M age = 11.1 years, SD age = 2.9 years). Of that sample, 1,074 were male and 154 were female. See Fischbach and Lord [2010] for additional information on SSC data collection, recruitment, diagnoses, and inclusion criteria.

This study was approved by The Children's Hospital of Philadelphia Institutional Review Board and adheres to the legal requirements of the United States.

Instruments

The DAS-II consists of 20 individually administered subtests divided into two overlapping batteries: Early Years (2:6 through 8:11 years) and School Age (7 through 17 years). This study analyzed only School Age data. The School Age core DAS-II battery consists of six subtests that contribute to three composite scores: verbal (word definitions and verbal similarities), nonverbal (matrices and sequential and quantitative reasoning), and spatial (pattern construction and recall of designs), as well as an overall composite derived from all six core subtests called general conceptual ability (GCA). Each composite score has a mean of 100 and a standard deviation of 15. The verbal, nonverbal, and spatial labels were retained for consistency with the first edition of the DAS but are assumed to measure crystallized intelligence (Gc), fluid reasoning (Gf), and visual–spatial ability (Gv), respectively, within the CHC (Cattell–Horn–Carroll) model of cognitive abilities [Elliott, Salerno, Dumont, & Willis, 2018; Schneider & McGrew, 2018]. The overall composite score (GCA) is assumed to measure *g*, the general factor of intelligence.

According to its author, the DAS-II was explicitly designed for the identification of cognitive strengths and weaknesses and has acquired considerable evidence supporting the reliability and validity of its scores [Elliott, 2007b; Elliott et al., 2018]. Independent test reviews have also supported the reliability and validity of DAS-II scores [Beran & Elliott, 2007; Marshall, McGoey, & Moschos, 2011].

Analyses

All analyses were conducted with version 15.1 of Stata. First, three cognitive difference scores were computed to quantify the cognitive profiles

- V-N = DAS-II Verbal score—nonverbal score
- V-S = DAS-II Verbal score—spatial score
- N-S = DAS-II Nonverbal score—spatial score

Second, a series of seven independent Welch's t-tests were conducted to evaluate the presence of mean differences between ASD and Norm groups on three DAS-II composite scores, three discrepancy scores, and one overall ability score [Welch, 1947]. Next, that series of Welch's t-tests were repeated for males and then for females. Third, a series of six dependent t-tests were conducted on the three composite and discrepancy scores in the ASD group to identify any mean differences within that group. Given a large number of statistical tests, significance was set at $P < 0.005$ to control the overall error rate. Finally, ROC analyses were conducted for each composite and difference score for the ASD and Norm groups and then separately for males and females. AUC values from each ROC analysis were computed using a nonparametric method given that Mardia's [1970] multivariate kurtosis was statistically significant at $P < 0.005$ [Youngstrom, 2014].

Results

Group Differences

As illustrated in Table 1, all DAS-II composite scores and the GCA were significantly ($P < 0.005$) lower for the ASD group than the DAS-II normative group. Of note, both the Norm and ASD samples included participants with a broad range of cognitive ability (GCA of 30–166 and 40–167, respectively). Additionally, the three discrepancy scores were significantly larger for the ASD group than the Norm group. Although statistically significant, standardized effect sizes were small (ranging from 0.14 to 0.42) using Cohen's [1988] guidelines. When separated by sex, group differences were more common for males (six of seven significant comparisons vs. three of seven significant comparisons for females). Within the ASD group, spatial composite scores were significantly higher than nonverbal composite scores and nonverbal composite scores were significantly higher than verbal composite scores although standardized effect sizes were small (d = 0.12–0.16). As expected for an autistic sample compared to a normative sample, there were significant sample sex differences.

Individual Differences

ROC analyses were conducted for the three DAS-II cluster scores and three discrepancy scores for the ASD group and the normative sample. Surprisingly, the highest AUC was observed not for a discrepancy profile, but for the simple nonverbal composite score (see Figure 2), which yielded an AUC value of 0.62, 95% CI [0.60–0.64], which would be considered low accuracy [Swets, 1996]. Thus, if two individuals were selected at random, one with ASD and one without ASD, the DAS-II nonverbal composite score would correctly identify the child with ASD 62% of

Table 1. Mean and Standard Deviation of DAS-II Index Scores for Males and Females in Normative and ASD Groups

	Norm			ASD			Δ		
	Males	Females	Total	Males	Females	Total	*P*	|*d*|	AUC
N	1,100	1,100	2,200	1,074	154	1,228	–	–	
Verbal	100.1 (15.4)	99.7 (14.9)	99.9 (15.2)	92.2 (22.5)[a]	93.4 (23.5)	92.4 (22.6)	<0.005	0.40	0.61 ± 0.01
Nonverbal	99.9 (15.3)	99.9 (14.3)	99.9 (14.8)	93.5 (18.6)[a]	88.6 (21.1)[a]	92.9 (19.0)	<0.005	0.42	0.62 ± 0.01
Spatial	100.3 (15.2)	99.0 (14.9)	99.7 (15.1)	95.4 (18.1)[a]	90.5 (19.0)[a]	94.8 (18.3)	<0.005	0.30	0.59 ± 0.01
GCA	100.2 (15.7)	99.5 (14.8)	99.9 (15.2)	94.1 (19.6)[a]	91.6 (21.8)[a]	93.8 (19.9)	<0.005	0.35	0.61 ± 0.01
V-N	0.22 (13.0)	−0.16 (12.6)	0.03 (12.8)	−2.66 (16.6)[a]	1.88 (15.3)	−2.12 (16.5)	<0.005	0.14	0.54 ± 0.01[b]
V-S	−0.21 (14.1)	0.66 (14.6)	0.22 (14.4)	−4.21 (20.0)[a]	1.36 (18.9)	−3.54 (20.0)	<0.005	0.21	0.56 ± 0.01
N-S	−0.46 (12.1)	0.81 (12.1)	0.18 (12.1)	−1.90 (14.3)	−1.86 (12.8)	−1.89 (14.1)	<0.005	0.15	0.54 ± 0.01
Age	12.5 (3.2)	12.5 (3.2)	12.5 (3.2)	11.0 (2.9)	11.5 (3.1)	11.1 (2.9)	<0.005	0.45	–

Abbreviations: V, verbal composite score; N, nonverbal composite score; S, spatial composite score; GCA, general cognitive ability score; AUC, area under the ROC curve (±1 standard error), ASD = sample of children with autism spectrum disorder from the Simons Simplex Collection.

[a]Same-sex comparisons between Norm and ASD group differences significant at $P < 0.005$.

[b]AUC values significantly different between males and females ($P < 0.005$).

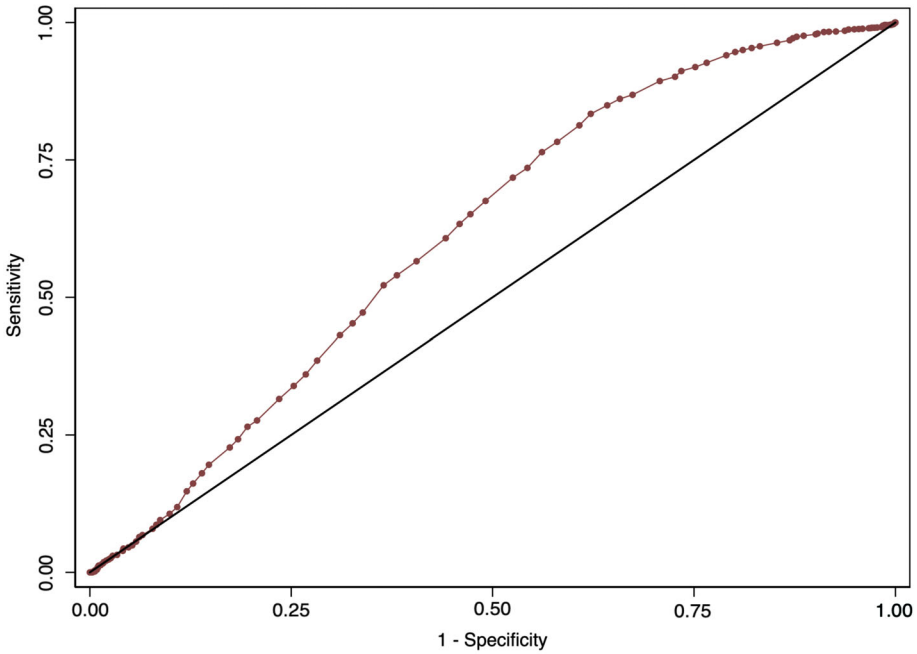

Figure 2. ROC analysis of nonverbal composite scores for children with autism spectrum disorder from the Simons Simplex Collection and the Differential Ability Scales—Second Edition School Age normative sample. Overall AUC = 0.62, 95% CI [0.60, 0.64].

the time [Kraemer et al., 2011]. AUC values changed very little (e.g., 0.63 vs. 0.62) when age-adjusted DAS-II scores were analyzed. For simplicity, non-adjusted DAS-II scores are subsequently reported. At the optimum cut score that minimized total classification error [Yovanoff & Squires, 2006], the sensitivity and specificity of the nonverbal composite score were 45.3% and 67.4%, respectively. The nonverbal composite score was the best discriminator for both males and females (AUC = 0.61, 95% CI [0.60–0.72] and 0.66, 95% CI [0.61–0.72], respectively), and no significant sex difference was observed in accuracy.

DAS-II discrepancy scores were less efficient in discriminating ASD and Norm groups than the DAS-II composite scores; significant differences were observed for the cognitive discrepancy scores V-N, V-S, and N-S comparisons with verbal, nonverbal, and spatial composite scores ($P < 0.005$). Overall, the V-N score produced an AUC of 0.54, 95% CI [0.52, 0.56]. However, the decision accuracy was better for males than females, with males obtaining an AUC value of 0.55, 95% CI [0.53–0.58] and females a value of 0.46, 95% CI [0.41–0.52]. This comparison is displayed in Figure 3.

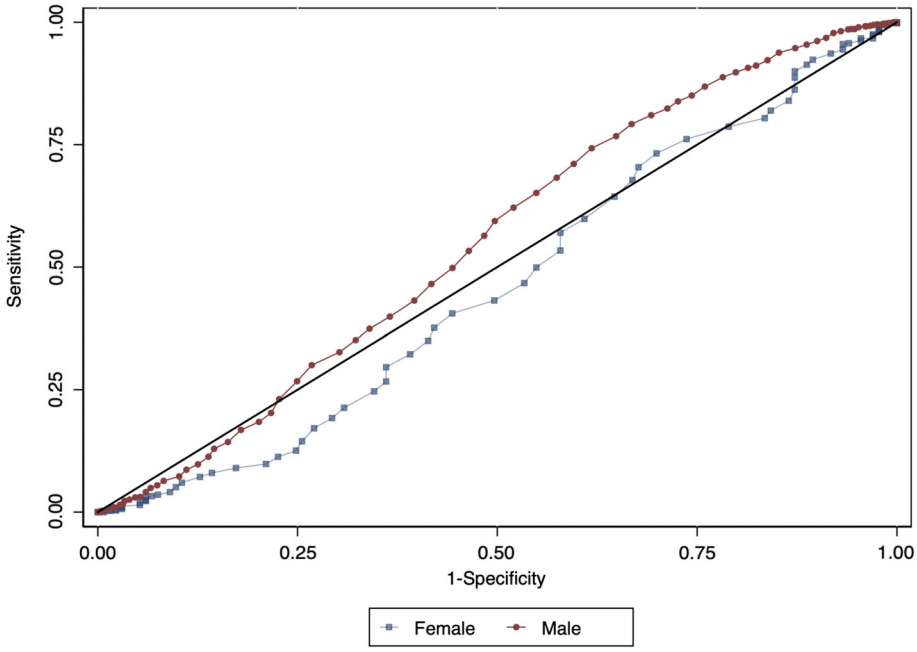

Figure 3. ROC analysis of verbal–nonverbal profile scores for male and female children with autism spectrum disorder from the Simons Simplex Collection and the Differential Ability Scales—Second Edition School Age normative sample. Overall AUC = 0.54, 95% CI [0.52, 0.56]. Male AUC = 0.55, 95% CI [0.53, 0.58] and female AUC = 0.46, 95% CI [0.41, 0.52].

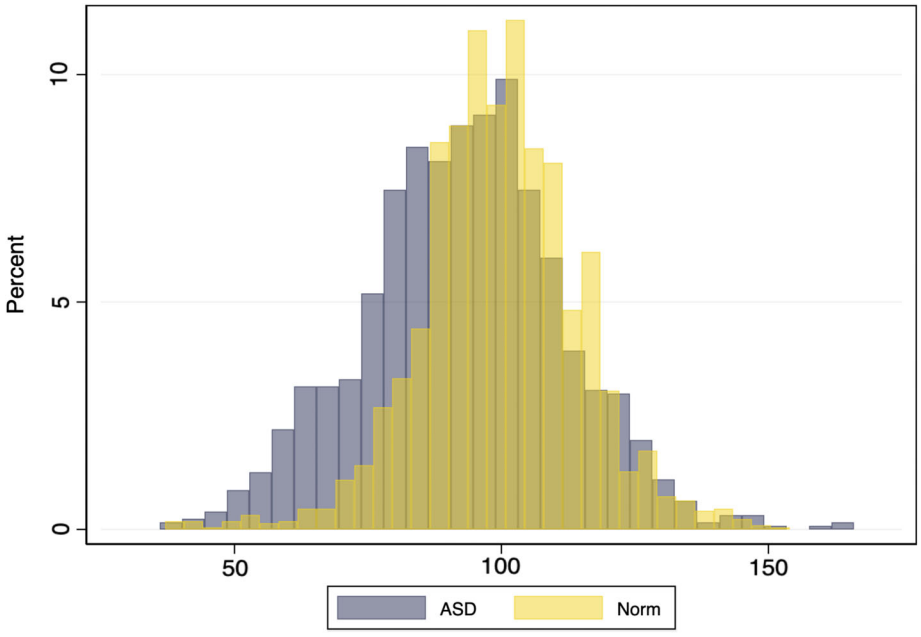

Figure 4. Distributional overlap of nonverbal composite scores for children with autism spectrum disorder from the Simons Simplex Collection (ASD) and the Differential Ability Scales—Second Edition School Age normative sample (Norm).

Norm group V-N discrepancy scores ranged from −49 to +42 whereas the ASD group discrepancy scores ranged from −68 to +52. As expected, the sensitivity of a V-N cut score of −68 was 100% whereas its specificity was 0%. Similarly, the sensitivity of a V-N cut score of +52 was 0% with a specificity of 100%. Sensitivity was 52% and sensitivity was 53% for an optimal V-N cut score of zero [Yovanoff & Squires, 2006]. Only four children with ASD exhibited V-N discrepancy scores lower than the Norm group bottom range and only five children with ASD

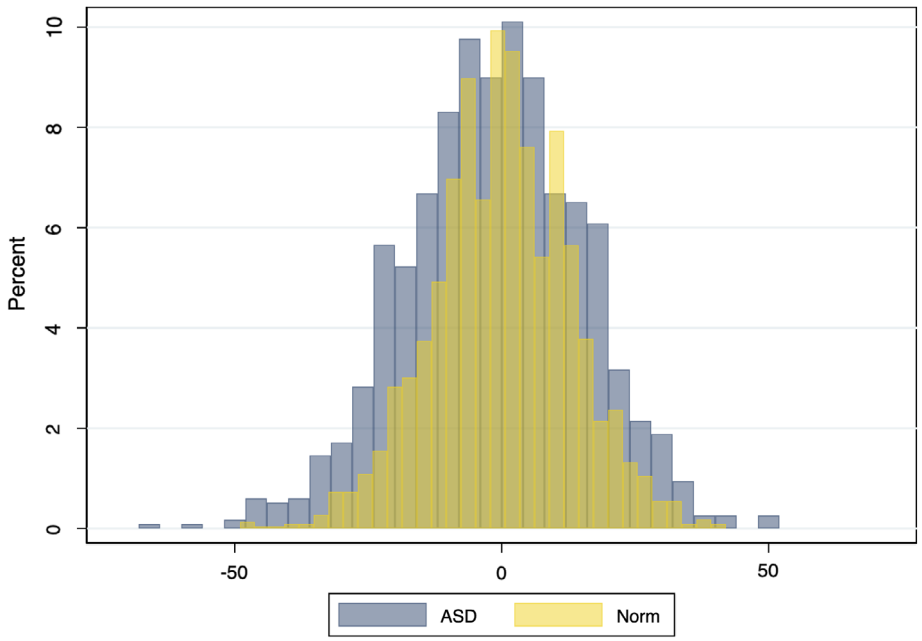

Figure 5. Distributional overlap of verbal–nonverbal profile scores for children with autism spectrum disorder from the Simons Simplex Collection (ASD) and the Differential Ability Scales—Second Edition School Age normative sample (Norm).

exhibited V-N discrepancy scores higher than the Norm group top range. Thus, only nine children with ASD (0.7%) obtained V-N discrepancy scores outside the Norm range. The extreme overlap in nonverbal composite score and V-N discrepancy score distributions is visually illustrated in Figures 4 and 5.

Discussion

The purpose of this study was to determine the degree to which DAS-II scores could accurately differentiate school-aged children with ASD from typically developing children using a large national sample of autistic children and the DAS-II standardization sample. It has often been hypothesized that uneven cognitive profiles with nonverbal scores higher than verbal scores [Klinger et al., 2018; Kroncke et al., 2016; Nader et al., 2015] are characteristic of autistic individuals. We replicated the finding of overall higher mean DAS-II nonverbal scores than verbal scores for autistic children. Further analysis focusing on the diagnostic predictive power at the individual level via ROC methods revealed that a verbal–nonverbal score discrepancy did not reliably distinguish autistic children from normative sample children (AUC = 0.54) and in fact, performed little better than chance. These sobering results may actually *over*estimate the clinical utility of a verbal–nonverbal score discrepancy because our comparison sample was predominantly typically developing, not children presenting for an autism diagnostic evaluation.

Based on group score differences, researchers have often concluded that cognitive profiles can be diagnostically useful [Chapman et al., 2011; Kroncke et al., 2016; Nader et al., 2015]. That is, accurate for making diagnostic decisions about individuals. However, test scores that "permit valid inferences about individuals will, of necessity, permit valid inferences about groups. The reverse, however, is not true" [Popham, 1993, p. 151]. In fact, to mistakenly conclude that differences at the group level also apply at the individual level is called the ecological fallacy [Robinson, 1950]. Research that focuses on correct decisions for individuals within groups is necessary to ascertain the clinical usefulness of cognitive profiles for diagnostic decision-making for an ASD diagnosis [Mandy et al., 2015; Meehl & Rosen, 1955].

ROC methods allow inferences about individuals [Kraemer et al., 2011; McFall & Treat, 1999; Swets, 1996; Youngstrom, 2014]. When ROC analyses were conducted with the Norm and ASD samples, individual decision accuracy for every discrepancy score was low and barely outperformed chance (AUC values of 0.54–0.62), inadequate for clinical decisions [Jarrett et al., 2018; Swets, 1996; Youngstrom, 2014]. These results make sense after noting the variance of the nationally representative, standardized normative sample itself: among children in the normative sample, V-N discrepancy scores ranged from −42 to +42. High discrimination and clinical utility would be difficult to achieve with the V-N discrepancy score, given this naturally occurring variability among all children.

High variability in cognitive ability at the group level has been frequently reported in association with ASD. We do not dispute this finding, and indeed observed larger standard deviations on all DAS-II composite and discrepancy scores for the ASD sample compared to the Norm sample. Previous research on the SSC and other samples with a group-based method described earlier (a priori cut score) noted a significantly higher proportion of autistic than normative children with V-N discrepancies greater than 16 points [Nowell et al., 2015]. Nowell et al. [2015] reported that 20% of autistic children in the SSC sample showed V < N score discrepancies of 16 or more points; however, 15% of the sample showed the opposite discrepancy (V > N), and 65% showed no discrepancy (V ≈ N), underscoring the difficulty of using the discrepancy to identify an individual with ASD. The Nowell et al. [2015] description of DAS-II group differences begs follow-up studies like ours that test ASD diagnostic prediction for individuals. In contrast to our DAS-II results, other measures show good discrimination: parent-completed screening questionnaires [e.g., SCQ, Corsello et al., 2007] have demonstrated AUC values with moderate accuracy (>0.75) and observational semi-structured surveys (e.g., ADOS [Brugha et al., 2012, Kamp-Becker et al., 2013]) have demonstrated AUCs with high accuracy (0.90) for autistic individuals.

Males and females with ASD exhibited different cognitive score patterns at the group level. In our sample, autistic males showed significantly higher nonverbal—but not verbal or spatial—scores than females (93.5 vs. 88.6). Verbal–nonverbal discrepancies showed significant differences between males and females, in opposing directions (V < N for males as expected, but V > N for females). The verbal–nonverbal discrepancy scores exhibited significantly better discriminatory effectiveness for males than females but did not reach a clinically useful level for either sex (AUC = 0.55 vs. 0.46, respectively). Based on mean sex differences, ASD appears to affect cognition differently in females and males in this sample; accordingly, we chose to present results separately for each sex so that a clinician can consider male-specific results when evaluating male patients and female-specific results when evaluating female patients, rather than one model with sex as a covariate that applies less readily to any individual patient.

The DAS-II verbal, nonverbal, and GCA composite scores were significantly more discriminative than discrepancy scores, suggesting that the meaning of patterns between scores is dwarfed by the meaning of overall lower average cognitive ability. The sensitivity and specificity of the nonverbal composite score were 45.3% and 67.4%, respectively, for an optimal cut score whereas sensitivity was 52% and specificity was 53%, respectively, for an optimal V-N discrepancy cut score. These sensitivities and specificities fall below the recommended threshold of 70% for developmental screening tests, and thus are inadequate for clinical decision-making [American Academy of Pediatrics, 2001]. It may surprise some that the DAS-II profiles were less clinically informative than the broader GCA, verbal, and nonverbal composite scores, but previous research demonstrates that the DAS-II is primarily detecting overall ability differences rather than any unique profile [Canivez & McGill, 2016; Dombrowski, Golay, McGill, & Canivez, 2018; Kotz, Watkins, & McDermott, 2008].

Cognitive profiles did not prove diagnostically useful, but might yield "subtypes" of ASD, particularly if such subtypes were created with additional information such as autistic symptoms (e.g., greater impairment in social approach, lower impairment in social motivation, presence or absence of a restricted interest, etc.) or comorbidities (e.g., presence of seizures, type of sleep disturbance, presence of gastrointestinal problems, etc.). Such subtypes could yield useful information about biological constructs, etiology, or prediction of treatment response. Future research might explore the utility of cognitive profiles in conjunction with symptom profiles for those purposes.

Limitations

Although male–female differences were found in this study, it is difficult to draw definitive conclusions due to the smaller number of females in the ASD sample than the Norm sample [Frazier et al., 2014]. This sample composition—possibly due to an ascertainment bias toward males, assessment biases, and other factors—might result in male-based inferences about autism. For a complete discussion of sex differences in the SSC sample and other large ASD samples, please see Frazier et al. [2014] and the accompanying commentary [Howe, Yatchmink, Viscidi, & Morrow, 2014]. Epidemiological studies typically find a 4:1 ratio of boys to girls, but the ratio of 7:1 boys vs. girls in this study may have allowed the subtle differences exhibited by girls to be overlooked due to reduced statistical power [Baio et al., 2018; Lerner et al., 2018]. Much needed efforts are underway to improve assessment and intervention for autistic girls [Klinger & Dudley, 2019].

Additionally, the SSC sample used in this analysis, while large and diverse in terms of race and ethnicity, is not representative of all US autistic children due to strict inclusion criteria. The SSC sample included only simplex ASD individuals, meaning individuals with no first-degree relatives with ASD, and excluded individuals with birth complications or with common medical or psychiatric disorders present in a first-degree relative. A previous study of toddlers suggests that simplex ASD presents with significantly lower cognitive ability than multiplex ASD, and modest group differences in verbal–nonverbal

discrepancies (small effect size), but no differences in autism severity [Dissanayake, Searles, Barbaro, Sadka, & Lawson, 2019]. Thus, these results may not generalize to the multiplex population. On the other hand, children were *not* excluded from the SSC sample based on the cognitive level, which is common in ASD research and a major strength of the sample highly relevant to the present study.

The size and direction of the group level verbal–nonverbal discrepancy have varied across samples in the literature, and the school-age SSC sample is another data point within the range. The variation likely reflects small sample sizes as well as sampling differences in functioning level [e.g., Siegel et al., 1996], co-occurring conditions [e.g., Kanai et al., 2017], the selection on IQ [e.g., Lennen et al., 2010], diagnostic criteria [e.g., Ghaziuddin & Mountain-Kimchi, 2004, Lincoln et al., 1998], and stringent genetic inclusion criteria [e.g., Nowell et al., 2015 and the present study]. All of these factors likely affect the verbal–nonverbal discrepancy in autistic individuals, providing a further rationale to avoid incorporating a discrepancy into diagnostic decision-making.

Our results are limited to school-age children (7–17 years). Although school-age children did not exhibit clinically useful cognitive profiles, future research could explore this question in preschool and adult samples. For example, Nowell et al. [2015] identified larger V-N discrepancies among autistic preschoolers in the SSC sample using the DAS-II Early Years.

Finally, we note that when evaluating sensitivity and specificity, the diagnoses of the two compared samples are of utmost importance. Ideally, the diagnostic utility of an autistic cognitive profile would be tested in a sample of children referred for a comprehensive ASD evaluation with cognitive testing; the discrepancy score AUC would be calculated by comparing clinically referred children who ultimately received an ASD diagnosis to those who did not. A study design with a clinically referred sample offers a more strenuous and ecologically valid test of the discrepancy score than our study design, which compared children with a clear diagnosis to normative children. A study with our design is likely to inflate artificially the AUC and overestimate the diagnostic utility of the tool, because a clinician's decision is usually *not* to classify the child as clearly ASD or clearly normative, as our study design implies. Rather, a clinician's decision is usually to classify the child as having ASD and/or other psychiatric or developmental disorder(s). Thus, the low sensitivity and specificity of the discrepancy scores reported in our study may actually *over*estimate their true clinical utility.

Conclusions

This study suggests that cognitive testing is not likely to be a valuable tool in the ASD diagnostic process in school-age children, even though cognitive testing remains an essential source of information to specify the level of functioning and intellectual disability. Further research on biological sex and gender identification differences in verbal–nonverbal discrepancies is needed.

Acknowledgments

The authors wish to thank Tiffany Ryan-D'Urso and Jen Swope for their assistance with this project. We are grateful to all of the families at the participating Simons Simplex Collection (SSC) sites, as well as the principal investigators (A. Beaudet, R. Bernier, J. Constantino, E. Cook, E. Fombonne, D. Geschwind, R. Goin-Kochel, E. Hanson, D. Grice, A. Klin, D. Ledbetter, C. Lord, C. Martin, D. Martin, R. Maxim, J. Miles, O. Ousley, K. Pelphrey, B. Peterson, J. Piggot, C. Saulnier, M. State, W. Stone, J. Sutcliffe, C. Walsh, Z. Warren, and E. Wijsman). We appreciate obtaining access to phenotypic data on SFARI Base. Approved researchers can obtain the Simons Searchlight population dataset described in this study (https://www.sfari.org/resource/simons-simplex-collection) by applying at https://base.sfari.org. Standardization data from the Differential Ability Scales-II (DAS-II). Copyright 1998, 2000, 2004, 2007 NCS Pearson, Inc. and Colin D. Elliott. Normative data copyright © 2007 NCS Pearson, Inc. used with permission. All rights reserved.

References

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, 67–88. https://doi.org/10.1080/15377903.2016.1236307

American Academy of Pediatrics. (2001). Developmental surveillance and screening of infants and young children. Pediatrics, 108, 192–196.

Ankenman, K., Elgin, J., Sullivan, K., Vincent, L., & Bernier, R. (2014). Nonverbal and verbal cognitive discrepancy profiles in autism spectrum disorders: Influence of age and gender. American Journal on Intellectual and Developmental Disabilities, 119, 84–99. https://doi.org/10.1352/1944-7558-119.1.84

Beran, T. N., & Elliott, C. D. (2007). Differential Ability Scales (2nd ed.). Canadian Journal of School Psychology, 22, 128–132. https://doi.org/10.1177/0829573507302967

Brugha, T. S., McManus, S., Smith, J., Scott, F. J., Meltzer, H., Purdon, S., … Bankart, J. (2012). Validating two survey methods for identifying cases of autism spectrum disorder among adults in the community. Psychological Medicine, 42, 647–656. https://doi.org/10.1017/S0033291711001292

Canivez, G. L., & McGill, R. J. (2016). Factor structure of the Differential Ability Scales–Second Edition: Exploratory and hierarchical factor analyses with the core subtests. Psychological Assessment, 28, 1475–1488.

Caterino, L. C. (2014). Cognitive, neuropsychological, academic, and adaptive functioning. In L. A. Wilkinson (Ed.), Autism spectrum disorder in children and adolescents: Evidence-based assessment and intervention in schools (pp. 75–99). Washington, DC: American Psychological Association.

Charman, T., Pickles, A., Simonoff, E., Chandler, S., Loucas, T., & Baird, G. (2011). IQ in children with autism spectrum disorders: Data from the Special Needs and Autism Project (SNAP). Psychological Medicine, 41, 619–627. https://doi.org/10.1017/S0033291710000991

Cohen, J. (1988). Statistical power analysis for the behavioral sciences. Hillsdale, NJ: Erlbaum.

Cormier, D. C., Wang, K., & Kennedy, K. E. (2016). Linguistic demands of the oral directions for administering the WISC-IV and WISC-V. Canadian Journal of School Psychology, 31, 290–304. https://doi.org/10.1177/0829573516643314

Corsello, C., Hus, V., Pickles, A., Risi, S., Cook, E. H., Leventhal, B. L., & Lord, C. (2007). Between a ROC and a hard place: Decision making and making decisions about using the SCQ. Journal of Child Psychology and Psychiatry, 48, 932–940. https://doi.org/10.1111/j.1469-7610.2007.01762.x

Dissanayake, C., Searles, J., Barbaro, J., Sadka, N., & Lawson, L. P. (2019). Cognitive and behavioral differences in toddlers with autism spectrum disorder from multiplex and simplex families. Autism Research, 12(4), 682–693. https://doi.org/10.1002/aur.2074

Dombrowski, S. C., Golay, P., McGill, R. J., & Canivez, G. L. (2018). Investigating the theoretical structure of the DAS-II core battery at school age using Bayesian structural equation modeling. Psychology in the Schools, 55, 190–207. https://doi.org/10.1002/pits.22096

Duvall, S. W., Huang-Storms, L., Hill, A. P., Myers, J., & Fombonne, E. (2020). No sex differences in cognitive ability in young children with autism spectrum disorder. Journal of Autism and Developmental Disorders, 50, 1770–1785. https://doi.org/10.1007/s10803-019-03933-1

Ehlers, S., Nydén, A., Gillberg, C., Sandberg, A. D., Dahlgren, S.-O., Hjelmquist, E., & Odén, A. (1997). Asperger syndrome, autism and attention disorders: A comparative study of the cognitive profiles of 120 children. Journal of Child Psychology and Psychiatry, 38, 207–217.

Elliott, C. D. (2007a). Differential ability scales (2nd ed.). San Antonio, TX: Harcourt Assessment.

Elliott, C. D. (2007b). Differential ability scales—Second Edition: Introductory and technical handbook. San Antonio, TX: Harcourt Assessment.

Elliott, C. D., Salerno, J. D., Dumont, R., & Willis, J. O. (2018). The differential ability scales–Second Edition. In D. P. Flanagan & E. M. McDonough (Eds.), Contemporary intellectual assessment: Theories, tests, and issues (4th ed., pp. 360–382). New York, NY: Guilford.

Elwood, R. W. (1993). Psychological tests and clinical discriminations: Beginning to address the base rate problem. Clinical Psychology Review, 13, 409–419.

Filipek, P. A., Accardo, P. J., Ashwal, S., Baranek, G. T., Cook, E. H., Dawson, G., … Volkmar, F. R. (2000). Practice parameter: Screening and diagnosis of autism: Report of the Quality Standards Subcommittee of the American Academy of Neurology and the Child Neurology Society. Neurology, 55, 468–479. https://doi.org/10.1212/WNL.55.4.468

Fischbach, G. D., & Lord, C. (2010). The Simons Simplex Collection: A resource for identification of autism genetic risk factors. Neuron, 68, 192–195.

Frazier, T. W., Georgiades, S., Bishop, S. L., & Hardan, A. Y. (2014). Behavioral and cognitive characteristics of females and males with autism in the Simons Simplex Collection. Journal of the American Academy of Child and Adolescent Psychiatry, 53, 329–340. https://doi.org/10.1016/j.jaac.2013.12.004

Ghaziuddin, M., & Mountain-Kimchi, K. (2004). Defining the intellectual profile of Asperger syndrome: Comparison with high-functioning autism. Journal of Autism and Developmental Disorders, 34, 279–284.

Giofrè, D., Provazza, S., Angione, D., Cini, A., Menazza, C., Oppi, F., & Cornoldi, C. (2019). The intellectual profile of children with autism spectrum disorder may be underestimated: A comparison between two different batteries in an Italian sample. Research in Developmental Disabilities, 90, 72–79. https://doi.org/10.1016/j.ridd.2019.04.009

Hausman-Kedem, M., Kosofsky, B. E., Ross, G., Yohay, K., Forrest, E., Dennin, M. H., … Ward, M. J. (2018). Accuracy of reported community diagnosis of autism spectrum disorder. Journal of Psychopathology and Behavioral Assessment, 40, 367–375. https://doi.org/10.1007/s10862-018-9642-1

Hayes, J., Ford, T., Rafeeque, H., & Russell, G. (2018). Clinical practice guidelines for diagnosis of autism spectrum disorder in adult and children in the UK: A narrative review. BMC Psychiatry, 18, 1–25. https://doi.org/10.1186/s12888-018-1800-1

Howe, Y. J., O'Rourke, J. A., Yatchmink, Y., Viscidi, E. W., Jones, R. N., & Morrow, E. M. (2015). Female autism phenotypes investigated at different levels of language and developmental abilities. Journal of Autism and Developmental Disorders, 45, 3537–3549. https://doi.org/10.1007/s10803-015-2501-y

Howe, Y. J., Yatchmink, Y., Viscidi, E. W., & Morrow, E. M. (2014). Ascertainment and gender in autism spectrum disorders. Journal of the American Academy of Child and Adolescent Psychiatry, 53(6), 698–700.

Huerta, M., & Lord, C. (2012). Diagnostic evaluation of autism spectrum disorders. Pediatric Clinics of North America, 59, 103–111. https://doi.org/10.1016/j.pcl.2011.10.01

Jarrett, M. A., Van Meter, A., Youngstrom, E. A., Hilton, D. C., & Ollendick, T. H. (2018). Evidence-based assessment of ADHD in youth using a receiver operating characteristic approach. Journal of Clinical Child & Adolescent Psychology, 47, 808–820. https://doi.org/10.1080/15374416.2016.1225502

Jones, R. M., & Lord, C. (2013). Diagnosing autism in neurobiological research studies. Behavioural Brain Research, 251, 113–124. https://doi.org/10.1016/j.bbr.2012.10.037

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, 807–821.

Kamp-Becker, I., Ghahreman, M., Heinzel-Gutenbrunner, M., Peters, M., Remschmidt, H., & Becker, K. (2013). Evaluation of the revised algorithm of Autism Diagnostic Observation Schedule (ADOS) in the diagnostic investigation of high-functioning children and adolescents with autism spectrum disorders. Autism, 17, 87–102. https://doi.org/10.1177/1362361311408932

Kanai, C., Hashimoto, R., Itahashi, T., Tani, M., Yamada, T., Ota, H., … Kato, N. (2017). Cognitive profiles of adults with high-functioning autism spectrum disorder with attention-deficit/hyperactivity disorder based on the WAIS-III. Research in Developmental Disabilities, 61, 108–115. https://doi.org/10.1016/j.ridd.2016.12.008

Klinger, L. G., & Dudley, K. M. (2019). Autism spectrum disorder. In M. J. Prinstein, E. A. Youngstrom, E. J. Mash, & R. A. Barkley (Eds.), Treatment of disorders in childhood and adolescence (4th ed., pp. 376–415). New York, NY: Guilford.

Klinger, L. G., Mussey, J. L., & O'Kelley, S. (2018). In S. Goldstein & S. Ozonoff (Eds.), Assessment of autism spectrum disorder (2nd ed., pp. 215–262). New York, NY: Guilford.

Kotz, K. M., Watkins, M. W., & McDermott, P. A. (2008). Validity of the general conceptual ability score from the Differential Ability Scales as a function of significant and rate interfactor variability. School Psychology Review, 37, 261–278.

Kraemer, H. C., Frank, E., & Kupfer, D. J. (2011). How to assess clinical impact of treatments on patients, rather than the statistical impact of treatments on measures. International Journal of Methods in Psychiatric Research, 20(2), 63–72. https://doi.org/10.1002/mpr.340

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

Kuriakose, S. (2014). Concurrent validity of the WISC-IV and DAS-II in children with autism spectrum disorder. Journal of Psychoeducational Assessment, 32, 283–294. https://doi.org/10.1177/0734282913511051

Kuschner, E. S., Bennetto, L., & Yost, K. (2007). Patterns of nonverbal cognitive functioning in young children with autism spectrum disorders. Journal of Autism and Developmental Disorders, 37, 795–807. https://doi.org/10.1007/s10803-006-0209-8

Lennen, D. T., Lamb, G. D., Dunagan, B. J., & Hall, T. A. (2010). Verbal prowess equals higher IQ: Implications for evaluating autism. Research in Autism Spectrum Disorders, 4, 95–101. https://doi.org/10.1016/j.rasd.2009.09.004

Lerner, M. D., Mazefsky, C. A., White, S. W., & McPartland, J. C. (2018). In J. N. Butcher, J. M. Hooley, & P. C. Kendall (Eds.), APA handbook of psychopathology: Child and adolescent psychopathology (Vol. 2, pp. 447–471). Washington, DC: APA.

Li, G., Du, Y., Luan, F., Li, M., & Ousley, O. (2017). IQ profiles and clinical symptoms of Chinese school-aged boys with autism spectrum disorder. European Journal of Psychiatry, 31 (2), 59–65. https://doi.org/10.1016/j.ejpsy.2017.04.001

Lincoln, A., Courchesne, E., Allen, M., Hanson, E., & Ene, M. (1998). Neurobiology of Asperger syndrome. In E. Schopler, G. B. Mesibov, & L. J. Kunce (Eds.), Asperger syndrome or high-functioning autism (pp. 145–163). New York, NY: Plenum.

Lord, C., Corsello, C., & Grzadzinski, R. (2014). Diagnostic instruments in autistic spectrum disorders. In F. R. Volkmar, S. J. Rogers, R. Paul, & A. Pelphrey (Eds.), Handbook of autism and pervasive developmental disorders (Vol. 2, 4th ed., pp. 609–660). Hoboken, NJ: Wiley.

Mandy, W., Murin, M., & Skuse, D. (2015). The cognitive profile in autism spectrum disorders. In M. Leboyer & P. Chaste (Eds.), Autism spectrum disorders: Phenotypes, mechanisms and treatments (pp. 34–45). Basel, Switzerland: Karger.

Mardia, K. V. (1970). Measures of multivariate skewness and kurtosis with applications. Biometrika, 57, 519–530.

Marshall, S., McGoey, K. E., & Moschos, S. (2011). Test review: Differential ability scales–Second Edition. Journal of Psychoeducational Assessment, 29, 89–93.

Matthews, N. L., Pollard, E., Ober-Reynolds, S., Kirwan, J., Malligo, A., & Smith, C. J. (2015). Revisiting cognitive and adaptive functioning in children and adolescents with autism spectrum disorder. Journal of Autism and Developmental Disorders, 45, 138–156. https://doi.org/10.1007/s10803-014-2200-0

Mayes, S. D., & Calhoun, S. L. (2008). WISC-IV and WIAT-II profiles in children with high-functioning autism. Journal of Autism and Developmental Disorders, 38, 428–439. https://doi.org/10.1007/s10803-007-0410-4

McDonnell, C. G., Bradley, C. C., Kanne, S. M., Lajonchere, C., Warren, Z., & Carpenter, L. A. (2019). When are we sure? Predictors of clinician certainty in the diagnosis of autism spectrum disorder. Journal of Autism and Developmental Disorders, 49, 1391–1401. https://doi.org/10.1007/s10803-018-3831-3

McFall, R. M., & Treat, T. A. (1999). Quantifying the information value of clinical assessments with signal detection theory. Annual Review of Psychology, 50, 215–241.

Mecca, T. P., Orsati, F. T., & de Macedo, E. C. (2014). Non-verbal cognitive profile of young children with autism spectrum disorders. Psychology, 5, 1404–1417. https://doi.org/10.4236/psych.2014.511151

Meehl, P. E., & Rosen, A. (1955). Antecedent probability and the efficiency of psychometric signs, patterns, or cutting scores. Psychological Bulletin, 52, 194–216.

Nader, A.-M., Jelenic, P., & Soulières, I. (2015). Discrepancy between WISC-III and WISC-IV cognitive profile in autism spectrum: What does it reveal about autistic cognition? PLoS One, 10(12), 1–16. https://doi.org/10.1371/journal.pone.0144645

Nowell, K. P., Schanding, G. T., Kanne, S. M., & Goin-Kochel, R. P. (2015). Cognitive profiles in youth with Autism Spectrum Disorder: An investigation of base rate discrepancies using the Differential Ability Scales—Second Edition. Journal of Autism and Developmental Disorders, 45, 1978–1988.

Ozonoff, S., Goodlin-Jones, B. L., & Solomon, M. (2005). Evidence-based assessment of autism spectrum disorders in children and adolescents. Journal of Clinical Child and Adolescent Psychology, 34, 523–540.

Popham, W. J. (1993). Educational evaluation (3rd ed.). Boston, MA: Allyn and Bacon.

Robinson, W. S. (1950). Ecological correlations and the behavior of individuals. American Sociological Review, 15, 351–357. https://doi.org/10.2307/2087176

Saulnier, C. A., & Ventola, P. E. (2012). Essentials of autism spectrum disorders evaluation and assessment. Hoboken, NJ: Wiley.

Scattone, D., Raggio, D. J., & May, W. (2012). Concurrent validity of the Leiter-R and KBIT-2 scales of nonverbal intelligence for children with autism and language impairments. Journal of Autism and Developmental Disorders, 42, 2486–2490. https://doi.org/10.1007/s10803-012-1495-y

Schneider, W. J., & McGrew, K. S. (2018). The Cattell-Horn-Carroll theory of cognitive abilities. In D. P. Flanagan & E. M. McDonough (Eds.), Contemporary intellectual assessment: Theories, tests, and issues (4th ed., pp. 73–163). New York, NY: Guilford.

Siegel, D. J., Minshew, N. J., & Goldstein, G. (1996). Wechsler IQ profiles in diagnosis of high-functioning autism. Journal of Autism and Developmental Disorders, 26, 389–406.

Styck, K. M., Aman, M. S., & Watkins, M. W. (2019). Diagnostic utility of the WISC-IV GAI > CPI cognitive score profile for a referred sample of children and adolescents with autism. Contemporary School Psychology, 23, 115–125. https://doi.org/10.1007/s40688-018-0172-3

Swets, J. A. (1996). Signal detection theory and ROC analysis in psychology and diagnosis: Collected papers. Mahwah, NJ: Erlbaum.

Taddei, S., & Contena, B. (2013). Cognitive performance in autism and Asperger's syndrome: What are the differences? Journal of Autism and Developmental Disorders, 43, 2977–2983. https://doi.org/10.1007/s10803-013-1828-5

Tager-Flusberg, H., & Joseph, R. M. (2004). Identifying neurocognitive phenotypes in autism. In U. Frith & E. Hill (Eds.), Autism: Mind and brain (pp. 43–66). New York, NY: Oxford University Press.

Volkmar, F., Siegel, M., Woodbury-Smith, M., King, B., McCracken, J., State, M., & The American Academy of Child and Adolescent Psychiatry Committee on Quality Issues. (2014). Practice parameter for the assessment and treatment of children and adolescents with autism spectrum disorder. Journal of the American Academy of Child & Adolescent Psychiatry, 53, 237–257.

Wechsler, D. (2011). Wechsler Abbreviated Scale of Intelligence—2nd Edition (WASI-II). San Antonio, TX: Psychological Corporation.

Wechsler, D. (2014). Wechsler Intelligence Scale for Children—5th Edition (WASI-V). San Antonio, TX: Psychological Corporation.

Weiner, I. B. (2003). Prediction and postdiction in clinical decision making. Clinical Psychology: Science and Practice, 10, 335–338. https://doi.org/10.1093/clipsy/bpg030

Welch, B. L. (1947). The generalization of students' problem when several different population variances are involved. Biometrika, 34, 28–35.

Youngstrom, E. A. (2014). A primer on receiver operating characteristic analysis and diagnostic efficiency statistics for pediatric psychology: We are ready to ROC. Journal of Pediatric Psychology, 39, 204–221. https://doi.org/10.1093/jpepsy/jst062

Yovanoff, P., & Squires, J. (2006). Determining cutoff scores on a developmental screening measure: Use of receiver operating characteristics and item response theory. Journal of Early Intervention, 29, 48–62.