

# Measurement of Nonverbal IQ in Autism Spectrum Disorder: Scores in Young Adulthood Compared to Early Childhood

Somer L. Bishop · Cristan Farmer ·  
Audrey Thurm

Published online: 20 September 2014  
© Springer Science+Business Media New York 2014

**Abstract** Nonverbal IQ (NVIQ) was examined in 84 individuals with autism spectrum disorder (ASD) followed from age 2 to 19. Most adults who scored in the range of intellectual disability also received scores below 70 as children, and the majority of adults with scores in the average range had scored in this range by age 3. However, within the lower ranges of ability, actual scores declined from age 2 to 19, likely due in part to limitations of appropriate tests. Use of Vineland-II daily living skills scores in place of NVIQ did not statistically improve the correspondence between age 2 and age 19 scores. Clinicians and researchers should use caution when making comparisons based on exact scores or specific ability ranges within or across individuals with ASD of different ages.

**Keywords** Cognitive ability · Intellectual disability · Adaptive behavior · Daily living skills

## Abbreviations

ASD Autism spectrum disorders  
ID Intellectual disability  
NVIQ Nonverbal IQ  
DLS Daily living skills

---

S. L. Bishop (✉)  
Center for Autism and the Developing Brain, Weill Cornell  
Medical College, 21 Bloomingdale Rd-Rogers Building,  
White Plains, NY 10605, USA  
e-mail: slb9013@med.cornell.edu

C. Farmer · A. Thurm  
Pediatrics and Developmental Neuroscience Branch, National  
Institute of Mental Health, National Institutes of Health, 10  
Center Drive, Bethesda, MD 20892, USA

## Introduction

Cognitive impairment in autism spectrum disorder (ASD) has long been studied, both as a common feature of individuals with the disorder and as a predictor of outcome. Studies of IQ in ASD at younger and older ages are greatly variable in the tests used and the types of scores reported. The overall distribution of IQ scores in ASD is skewed, with epidemiological studies indicating widely varying rates of comorbid intellectual disability (ID) depending on methodology (e.g., demographics of participants, measures used, inclusion/exclusion criteria) and location. In 2012, the Centers for Disease Control reported prevalence estimates of ID in ASD by US state ranging from 13 to 54 % (CDC 2012), while a recent review of 15 studies found the median prevalence estimate of ID in ASD to be 65 % (Dykens and Lense 2012). Outcome studies have shown that IQ in childhood predicts outcome, with some studies demarcating an IQ of 50 as critical for functional outcomes (Gillberg and Steffenburg 1987) and others using a categorical cut point of 70 (Howlin et al. 2004).

Unfortunately, while the importance of IQ to ASD clinical practice and research is widely recognized, detailed information about IQ in adults with ASD and how it relates to childhood IQ is limited. There has also been little discussion of the measurement issues related to testing individuals at multiple points in development. A recent review found that only five studies reported longitudinal IQ data from individuals below the age of 5 years through at least 18 years (Magiati et al. 2014). Furthermore, this literature has been biased by sample selection that may affect generalizability. For example, in one of the largest longitudinal studies to assess IQ systematically, Howlin et al. (2004) reported stability in general nonverbal IQ (NVIQ) ranges (i.e., >100, 70–99, and 50–69) for 68 individuals followed

from approximately age 7 years through an average age of 29. However, only individuals with NVIQ of at least 50 prior to age 16 were included in the analyses. Howlin et al. (2014) recently published a second follow-up of individuals from this sample, this time restricting analyses to 60 participants with childhood IQ greater than 70. Again, IQs in this sample were generally stable for those who could complete a standardized test, but one-quarter of the participants were excluded from some analyses because they could not complete these tests as adults, and for other analyses, a “best-estimate” IQ derived from the Vineland Adaptive Behavior Scales was substituted for IQ (Howlin et al. 2014).

The issue of cognitively non-representative adult ASD samples is widespread. A survey of articles published in an autism journal indicated that of the studies that reported IQ in adult participants with autism, the majority (77 %) required an IQ of at least 75 (ranging up to 90) for inclusion. Just three of 30 studies included participants with “low” IQ (presumed to be less than 70) (Dykens and Lense 2012). This may be partially due to the fact that studies of adults with lower cognitive abilities are severely limited by the lack of standardized intelligence tests for these individuals. Because many adults with ASD are unable to achieve a basal score on any test that allows for comparison with normative samples (e.g., a “floor” effect), some researchers exclude adults who cannot obtain valid scores, whereas others assign the lowest possible standard score on the test. Another alternative is to use age equivalents and calculated ratio scores from tests designed for children. The practice of deriving ratio IQs is common for individuals with ASD (Munson et al. 2008), but use of ratio IQ in adults and in longitudinal research is limited by artificial deflation of scores by an increasing denominator (chronological age) (Aiken 1996).

On the other hand, it is not clear how best to counteract the problems associated with the use of ratio IQ. One could create alternative standardized scores derived from raw scores below the “floor” for frequently used IQ tests (Hessl et al. 2009), but many individuals with low cognitive abilities are unable to achieve any points on tests with a lower age limit around school-age. Some researchers have used standard and ratio scores interchangeably (Sigman and McGovern 2005; Turner et al. 2006), including using standard scores from measures such as the Vineland Adaptive Behavior Scales instead of traditionally defined cognitive tests (Coplan and Jawad 2005; Howlin et al. 2014). This practice has little empirical support, particularly in older samples, although it is compelling to consider adaptive behavior given its relevance to the diagnosis of ID.

The lack of consensus about how best to measure IQ in adults with ASD is important, because IQ is a critical

variable in ASD research that has major implications regarding treatment outcomes (Eldevik et al. 2010) and genetic etiology of the disorder (Girirajan et al. 2013). As increasing numbers of individuals with ASD move into adulthood, there is a clear need to understand more about the measurement and meaning of cognitive scores in adults. In particular, a better understanding of ratio IQ is necessary for both clinicians and researchers to accurately describe, recognize, and treat adults with ASD. For clinicians, it is important to know what an IQ score means about an individual’s trajectory and prognosis. This information is also necessary for researchers in order to understand the limitations of data from studies reporting on within-subject correspondence between early and late IQ scores and ID categories, as well as to inform how individuals in large cross-sectional datasets can be grouped by IQ (e.g., for phenotyping and genotyping studies).

The current study was conducted to address the question of correspondence between cognitive ability in early childhood and adulthood. Though previous studies have established that early IQ is the single best predictor of adult independence, the extent to which actual scores in very young children are related to scores at older ages is not well understood. We examined these questions in a sample of individuals with ASD who were first assessed at age 2 and followed into young adulthood. Consistent with the other published longitudinal studies of IQ in ASD, analyses focused on assessing the stability of NVIQ scores, because there is often a large split in IQ scores, with verbal (and therefore full-scale) scores more influenced by language deficits (Joseph et al. 2002). Specific questions were: (1) To what extent do nonverbal cognitive scores and ID classification of individuals with ASD at age 19 correspond to scores and ID classification obtained at ages 2 and 3? (2) Does this differ based on the individual’s cognitive level?

## Methods

### Participants

Participants were 84 individuals with ASD, 87 % ( $n = 73$ ) male and 88 % ( $n = 74$ ) Caucasian, followed in an ongoing longitudinal study on the early diagnosis of autism (Lord et al. 2006). The original cohort included 192 children referred for possible autism and 22 children with non-spectrum developmental delays recruited as controls, who were first assessed around age 2 and re-assessed at the approximate ages of 3, 5, 9, and 19 years (see Table 1 for demographics). In the larger study, not all individuals participated at every time point and some individuals were lost to follow-up during the course of the study. Individuals were only eligible for inclusion in the current analyses if

**Table 1** Cognitive and ADOS test information by visit,  $N = 84$ 

	Age 2	Age 3	Age 19
Age, $M \pm SD$	28.9 $\pm$ 5.2 m	41.8 $\pm$ 5.8 m	19.0 $\pm$ 1.0 y
NVIQ			
$M \pm SD$	66.3 $\pm$ 20.6	63.5 $\pm$ 21.1	53.8 $\pm$ 38.8
Range	(13,109)	(17,110)	(3,128)
Ratio NVIQ, $n$ (%)	84 (100)	84 (100)	58 (67)
Test used, $n$ (%)			
Mullen	84 (100)	82 (98)	19 (21)
DAS	0 (0)	2 (2)	38 (45)
WASI	0 (0)	0 (0)	28 (33)
ADOS module, $n$ (%)			
PLADOS	84 (100)	77 (92)	0 (0)
Module 1	0 (0)	3 (4)	33 (40) <sup>a</sup>
Module 2	0 (0)	2 (2)	22 (26)
Module 3	0 (0)	0 (0)	1 (1)
Module 4	0 (0)	0 (0)	28 (33)

Sample size differed for age 3 ADOS ( $n = 82$ ). At the age 19 visit, one individual was 213 months of age and therefore within the age range to receive a standard score on the DAS

NVIQ nonverbal intelligence quotient, DAS differential abilities scales (preschool or school age), WASI Wechsler abbreviated scale of intelligence

<sup>a</sup> At age 19,  $n = 30$  (36 %) of children had single words or no words according to the ADOS overall level of non-echoed language code

they received cognitive assessments at ages 2, 3, and 19 and exhibited a clear history of ASD from childhood, defined as two or more best-estimate clinical diagnoses of ASD (described below) at ages 2, 9, and 19 (85 % received an ASD diagnosis at all time points).

## Measures

### *Mullen Scales of Early Learning (MSEL; Mullen 1995)*

The MSEL is a standardized developmental test for children birth to 5 years, 8 months. Subscales yield T-scores and age equivalents and can be combined to provide an estimate of overall developmental functioning, the Early Learning Composite, with standard scores ranging from 49 to 155. The MSEL has good internal, test–retest, and interrater reliabilities (Mullen 1995).

### *Differential Ability Scales (DAS; Elliot 1990) and Differential Ability Scales, 2nd Edition (DAS-II; Elliott 2007)*

The DAS and DAS-II are tests of cognitive ability that are standardized for children aged of 2 years, 6 months, to 17 years, 11 months. Both versions of the DAS have good internal, test–retest, and interrater reliabilities, and the

content of the two versions is similar enough to warrant combination for analyses. Standard scores for NVIQ range from 45 to 165 (with extended norms down to 25 in the DAS-II). In a sample of young children with and without ASD (some of whom were included in the current study), the DAS and MSEL showed good convergent validity (Bishop et al. 2011).

### *Wechsler Abbreviated Scale of Intelligence (WASI; Wechsler 1999)*

The WASI is a test of cognitive ability that is appropriate for individuals aged 6–89 years. Two of the four subtests contribute to performance IQ (NVIQ), with standard scores ranging from 53 to 157.

### *Vineland Adaptive Behavior Scales, Second Edition (Vineland-II; Sparrow et al. 2005)*

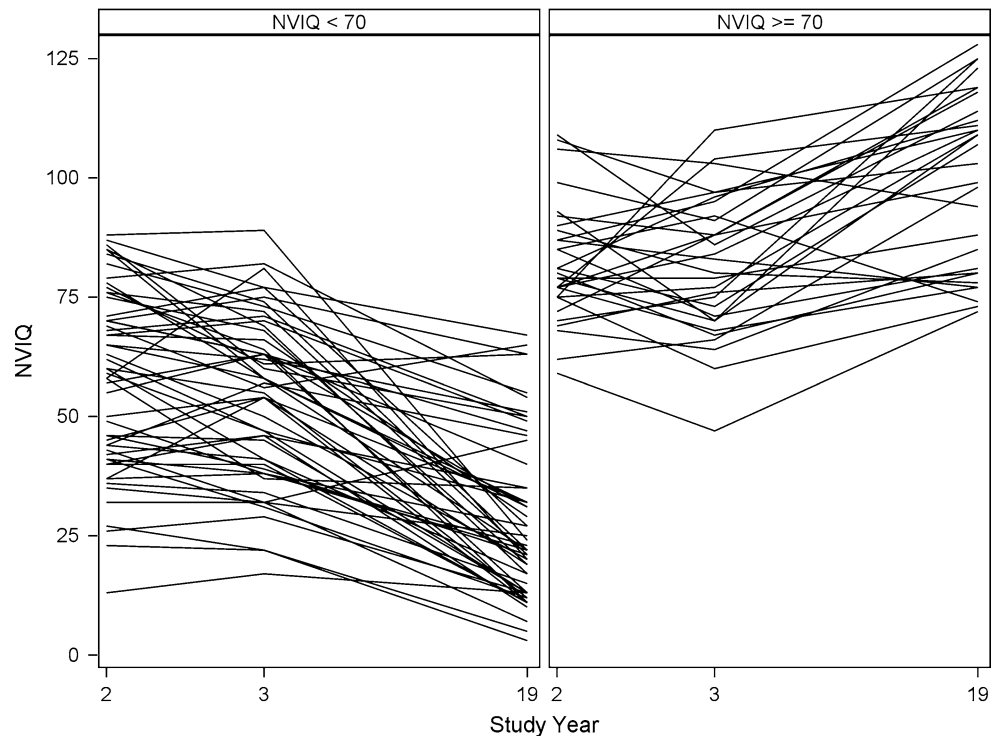
The Vineland-II is a semi-structured caregiver interview that produces standard scores in several domains. The current study utilized daily living skills (DLS) as a measure of adaptive functioning, because it is thought to be less influenced by ASD symptoms than the other domain scores, particularly in individuals with significant cognitive impairment (Duncan and Bishop 2013; Kraijer 2000).

## Procedures

All procedures were approved by the Institutional Review Board, and all participants provided informed consent. The assessment protocol included administration of the Autism Diagnostic Interview-Revised (ADI-R; Le Couteur et al. 2003) and the Vineland-II (or the first edition) to parents. Children completed the Pre-linguistic Autism Diagnostic Observation Schedule (PL-ADOS) (DiLavore et al. 1995) at ages 2 and 3, and at age 19, completed either Module 3 or 4 of the standard ADOS (Lord et al. 2000) or the Adapted version of Modules 1 or 2 (Hus & Lord, personal communication). A best estimate clinical diagnosis of ASD or non-ASD was assigned based on all information gathered during each assessment (Lord et al. 2006).

Table 1 describes IQ test administration by time point. The MSEL was administered to all children at age 2, and the DAS was used if possible at age 3. At age 19, the WASI was administered if standard scores could be achieved, followed in preference by the DAS-II and the MSEL. Because participants at age 19 were all outside of the standard age range for administration of the MSEL, ratio NVIQ scores were calculated (average age equivalent of the Visual Reception and Fine Motor subscales divided by

**Fig. 1** IQ across study years by NVIQ broad categorization at age 19 (n = 84)



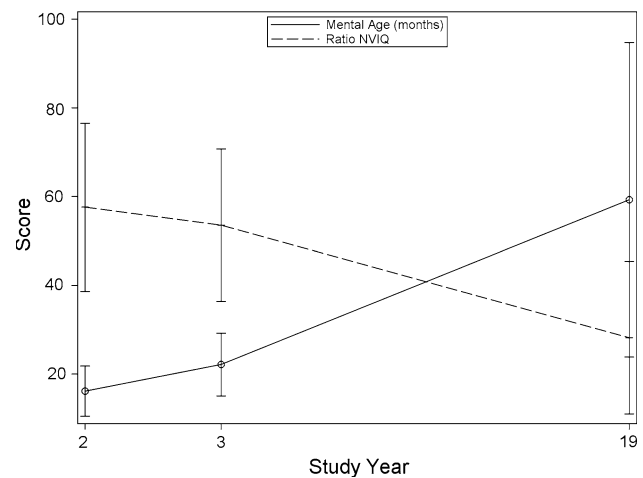
chronological age, multiplied by 100; Bishop et al. 2011). In all but one case, age-based norms were not available for participants who received the DAS-II at the age 19 visit. DAS-II ratio NVIQ scores at the age 19 time point were calculated using a chronological age denominator of exactly 18 years, regardless of actual age.

**Data Analysis**

SPSS Version 21.0 was used for all analyses. All standard linear regression and stability (Cohen’s kappa) analyses in the current study were exploratory. As partial correction for multiple comparisons, alpha was set to .01.

**Results**

Cognitive scores by study time point are reported in Table 1. Average NVIQ was lower at age 19 than at age 2. For 67 % (n = 56) of the sample at age 19, ratio IQs were derived from tests standardized for children. Categorization of IQ at 70 was a good proxy for ratio versus standard score (one individual had a ratio score over 70 and three individuals had standard scores less than 70). Figure 1 shows change in NVIQ over time based on this grouping. Individuals with NVIQ above 70 at age 19 showed a trend of increasing scores over time, while those with NVIQ scores



**Fig. 2** Nonverbal mental age versus ratio score in participants with low NVIQ at study year 19 (n = 54). Note n = 54 participants with NVIQ score <70 at the age 19 time point. One participant received standard score; remainder received ratio score at age 19 time point

below 70 at age 19 showed a trend of decreasing ratio NVIQ over time. To better understand the declining ratio NVIQ, a plot of mental age over time is shown in Fig. 2. These results show that declining average ratio NVIQ in the individuals with NVIQ below 70 at age 19 is due not to an actual decline in cognitive ability, but rather to a slower-than-expected gain in skills over time.

**Table 2** Age 2 and age 19 cognitive/adaptive scores by category at age 19

	Age 19 NVIQ		Age 19 NVIQ	
	≥70	<70	≥50	<50
<i>n</i>	30	54	39	45
Age 2 NVIQ	82.10 ± 12.43	57.57 ± 18.94	80.67 ± 11.79	53.91 ± 19.37
≥70, n (%)	26 (31)	17 (20)	–	–
<70, n (%)	4 (5)	37 (44)	–	–
NVIQ nonverbal developmental quotient/nonverbal intelligence				
≥50, n (%)	–	–	39 (46)	24 (29)
<50, n (%)	–	–	0 (0)	21 (25)
Age 2 VABS DLS	71.93 ± 8.31	66.58 ± 7.73 <sup>a</sup>	71.08 ± 7.69 <sup>a</sup>	66.25 ± 8.25
Age 19 NVIQ	99.93 ± 19.38	28.17 ± 17.19	90.15 ± 24.39	22.29 ± 11.61
Age 19 VABS DLS	79.00 ± 13.15 <sup>a</sup>	48.98 ± 14.29	75.13 ± 14.14 <sup>a</sup>	46.24 ± 13.55

*VABS DLS* Vineland-II daily living skills  
<sup>a</sup> VABS DLS data missing for one participant in this group

**Stability of NVIQ Scores**

Age 2 NVIQ shared 44 % of variance with NVIQ at age 19 ( $r = .67, p < .001$ ). This relationship was driven by individuals with low NVIQ scores (<70) at age 19 ( $R^2 = .25$  for NVIQ < 70,  $R^2 = .05$  for NVIQ ≥ 70). NVIQ at age 3 was similarly associated with age 19 NVIQ ( $r = .74, p < .001$ ), sharing slightly more of the variance than age 2 scores ( $R^2 = .55$ ). The amount of variance explained by age 3 scores was similar between groups based on age 19 NVIQ above or below 70.

**Stability of Broad Score Classifications**

Though NVIQ for many individuals remained stable between ages 2 and 19, a significant minority of individuals showed large changes (Fig. 1). To better understand the relationship between early and later cognitive scores, we assigned participants to broad categories of cognitive ability (cut points at 50 and 70).

The stability of these categorizations across time points is shown in Table 2. The proportion of individuals above the 50 and 70 cutoffs decreased between age 2 (75 and 51 %, respectively) and age 19 (46 and 36 %). Only half ( $n = 21$  of 45) of the participants with NVIQ < 50 at 19 had been classified as such at age 2. For dichotomization at 50, Cohen’s kappa was in the moderate range at age 2 ( $K = .49$ ) and at age 3 ( $K = .45$ ). Categorization by 70 was similarly stable for those with lower NVIQ, but 40 % of participants with NVIQ above 70 at age 2 moved into the <70 category at age 19 (Cohen’s kappa age 2  $K = .50$ ; age 3  $K = .58$ ). To further illustrate the potential clinical consequences of this downward movement, Table 3 shows classification changes into more severe ranges of DSM-IV-TR defined ID.

**Table 3** Comparison of age 2 and age 3 with age 19 NVIQ intellectual disability categories

	Age 19 NVIQ (%)						
	Avg.	Border.	Mild	Mod.	Sev.	Pro.	Total
<b>Age 2</b>							
Average	47	24	12	0	12	6	20
Borderline	42	12	15	4	19	8	31
Mild	10	10	15	15	25	25	24
Moderate	0	0	0	19	38	44	19
Severe	0	0	0	0	0	100	5
Profound	0	0	0	0	0	100	1
Total	25	11	11	8	21	24	100
<b>Age 3</b>							
Average	86	7	0	0	7	0	17
Borderline	33	19	24	5	19	0	25
Mild	8	12	15	12	27	27	31
Moderate	0	7	0	14	36	43	17
Severe	0	0	0	13	13	75	10
Profound	0	0	0	0	0	100	1
Total	25	11	11	8	21	24	100

Categories correspond to DSM-IV-TR. Average: ≥85, borderline: 70–84, mild: 50–69, moderate: 35–49, severe: 20–34, profound: <20

**Relationship Between Early NVIQ and Later Adaptive Behavior**

The average NVIQ score in the low broad categories at age 19 (i.e., scores <50 and <70) was lower than their average NVIQ scores at age 2 (Table 2). Based on the rise in age equivalents (Fig. 2), we explored whether this was also reflected in their level of adaptive behavior. We substituted standardized scores from the Vineland-II DLS as an outcome at age 19, both as a continuous variable and classified into the broad categories. Table 2 shows the age 19 DLS

scores in each of the broad categories; for individuals with  $NVIQ < 70$ , the average difference between NVIQ and DLS was  $20.82 \pm 10.96$ , with DLS scores significantly higher than NVIQ (paired  $t$  test =  $-13.96$ ,  $p < .001$ ). For those with  $NVIQ < 50$ , DLS scores were an average of  $23.96 \pm 7.97$  higher than NVIQ ( $t = -20.17$ ,  $p < .001$ ). For cognitively higher functioning individuals, DLS scores were significantly lower than NVIQ scores ( $\geq 70$ :  $20.28 \pm 14.45$ ,  $t = 7.55$ ,  $p < .001$ ;  $\geq 50$ :  $14.26 \pm 17.39$ ,  $t = 5.06$ ,  $p < .001$ ). Although the average DLS standard scores at age 19 were closer to the age 2 NVIQ than were the average age 19 ratio NVIQs, age 19 DLS standard scores were not better predicted by NVIQ at ages 2 and 3 than was NVIQ at age 19. NVIQ scores at the age 2 and age 3 time points both shared 41 % of variance with age 19 Vineland-II DLS scores (age 2:  $r = .64$ ,  $p < .001$ ; age 3:  $r = .64$ ,  $p < .001$ ). These relationships were stronger in participants with age 19  $NVIQ < 70$  (age 2:  $R^2 = .19$  vs.  $R^2 = .11$ ; age 3:  $R^2 = .16$  vs.  $R^2 = .04$ ).

## Discussion

In this study, we observed a trend of declining NVIQ in individuals with ASD between toddlerhood and young adulthood. This downward trend was primarily evident in individuals who received scores below 70. The magnitude of these declines was large enough that it would have affected DSM-IV-TR mental retardation (MR; now termed ID in DSM-5) classifications assigned on the basis of NVIQ; about 15 % of the sample moved from above 70 at age 2 to below 70 at age 19, and about 40 % of the entire sample moved into the severe and profound categories. Thus, although the NVIQ of many participants (especially those with higher NVIQ) was stable or increased over time, the mean NVIQ in this sample declined substantially between ages 2 and 19.

As should be expected, broad IQ categorizations were more stable than specific DSM-IV-TR MR categories over time. Clinically, these results have implications for the concept of delay versus deviance in young children with ASD. The baseline assessment in this study occurred at an age when low scores are clinically conceived of as delays (i.e., global developmental delay) rather than deviance (i.e., intellectual disability). However, the results of this study suggest that the concept of delay may not be meaningful for children with ASD with the lowest scores, as children with NVIQ in the ID range by age 3 were unlikely to move out of that range by age 19. Similarly, individuals who scored in the average range or above by age 3 tended to receive scores in that range at age 19, although there was inter-individual variability. The greatest variability was observed for individuals in the borderline range: only 11 %

retained this designation at age 19 (about half moved into the average category; the remainder moved into the ID range). Based on these findings, professionals should communicate to parents that many young children with ASD who exhibit significant nonverbal cognitive “delays” remain in the ID range as adults. On the other hand, they should also be careful to not provide parents of preschoolers with guarantees that their children will remain outside of the ID range based on average or borderline NVIQ scores.

Decline in NVIQ from toddlerhood to adulthood was most common in individuals with IQs below 70 at age 19. Most of these individuals could not complete enough items to receive standardized scores on age appropriate tests, so they were administered tests designed for children, and ratio IQs were calculated from the age equivalents provided by these tests. The observed decline in IQ scores among these individuals therefore highlights the potential limitations of using calculated ratio IQs at older ages to study changes in an individual over time. Within an individual, comparison of later ratio scores with prior IQ scores is invalid after a certain age that remains undetermined, but is likely well before 18 years (Aiken 1996). This is because the artificial deflation of ratio scores by an ever-increasing denominator in later school age and adolescence causes the false appearance that ability decreases over time, when in reality the trend of decreasing NVIQ is most likely due to measurement limitations associated with using age-inappropriate tests.

Use of ratio scores in older individuals also results in a compressed range at the lower end of the IQ continuum, such that differences in skills and abilities are not adequately reflected by the scale of ratio scores. For example, using a chronological age denominator of 18 years, a young adult with a nonverbal mental age of 48 months would receive a ratio score of 22, and a young adult with a nonverbal mental age of 24 months would receive a ratio score of 11. Although functioning at the level of an average 2-year-old is very different from functioning at the average level of a 4-year-old, these scores are similar when interpreted using the conventions imported from IQ (i.e., the scores are within one standard deviation and both in the profound range of ID). These scores exemplify the wide variability in functioning that exists in the DSM profound range of ID. Furthermore, this problem of a “shrinking” IQ range has implications for research design and interpretation, in that it creates an illusion of bi-modally distributed nonverbal abilities in adolescents/young adults with ASD. The alternative of using age equivalents (mental ages) with age-appropriate standardized instruments is also not possible, due to the dilemma of either not appropriately controlling for age of the individual (when tests standardized on younger ages are used) or limited psychometrics

available in IQ tests standardized for adults. According to data presented in this study, the majority of those who can achieve standard scores on IQ tests will score in or above the average range of ability, and most of the remainder will score in the moderate, severe, or profound range of ID. Commonly used intelligence scales for adults simply do not capture the variability among adults in this lower range, illustrated most emphatically by the lack of standardization below scores of 40–50. Recently published IQ tests for both preschool children (Roid 2003) and for nonverbal individuals (Wechsler and Naglieri 2006) with standardization extended to IQs of 10 may be helpful in this regard.

Our results suggest that use of standard scores from the Vineland-II or another adaptive behavior measure may not be a useful substitute for IQ at older ages, if trajectory of IQ is the goal. The Vineland-II daily living skills score at age 19 was not more strongly predicted by early NVIQ scores than was NVIQ at age 19. On the other hand, since Vineland-II daily living scores did not decline as precipitously as IQ scores and were generally higher than IQ scores in individuals in the ID range at age 19, measurement of adaptive behavior that directly relates to needed supports will provide more useful clinical information about individuals with more severe cognitive impairment, especially for those who cannot complete tests standardized for their age-range. This is consistent with general clinical practice and with evidence from previous longitudinal studies of ID regarding age-related changes (Chadwick et al. 2005) that support newer definitions of intellectual disability which include a broader focus on adaptive behavior (American Psychiatric Association 2013; Greenspan and Woods 2014; Schalock et al. 2011).

Quantification of IQ across the full range of functioning in adults with ASD remains a major challenge for the field. This is true for cross-sectional analysis, but even more so for longitudinal and treatment studies evaluating trajectories. There is extraordinary variability in the tests used to estimate cognitive functioning in ASD, and the heterogeneity in test selection necessarily increases with age. The use of chronological age-inappropriate tests in lower functioning individuals with ASD is informative, but interpretation of the resulting scores is complex. Since scores between these tests and age-appropriate tests are not directly comparable, classifications may be a better approach to comparing cognitive functioning of adolescents/adults. However, with respect to cross-sectional analyses, it is imperative to consider that the abilities that correspond to borderline or mild ID in a young child, for example, may correspond to severe or profound ID in an older child. Similarly, young children who score in the severe or profound range of intellectual disability are likely to be quite different than individuals who score in these ranges at older ages. Furthermore, adolescents or adults excluded from research

based on a minimum NVIQ requirement (e.g., 50) may have been eligible when they were younger, leading to adult samples that are likely to be even more cognitively skewed than child samples. Thus, researchers must be careful to remember that although the construct of intelligence is theoretically static, IQ scores and classifications are not (Burgaleta et al. 2014).

### Limitations

The most significant limitation of this study was attrition, such that less than half of the ASD cohort enrolled at the age 2 time point had cognitive data at the age 19 time point. Other studies have found that based on prior testing, those considered “lower functioning” are more difficult to retain (Eaves and Ho 2008; Gillespie-Lynch et al. 2012). The sample reported upon in this study did not differ significantly from the remainder of the sample diagnosed with ASD on age 2 demographic or phenotypic variables. Still, it is impossible to know whether individuals who were lost-to-follow-up would have differed at age 19. In addition, although the limited availability of appropriate tests for individuals with ASD and ID necessitated the use of several different cognitive measures, this variability is a limitation. At the age 19 time point, clinicians were advised to first attempt the WASI, followed by the DAS-II and then the MSEL, but decisions about which test to start with were based on clinical judgment that could have varied between clinicians. There are no reports on convergent validity of IQ tests when they are administered outside of the standard age range, and it is possible that individuals in the current study who were administered the MSEL may have scored differently had they been administered the DAS-II, for instance. Some tests with wider age ranges (e.g., the Stanford-Binet) may have provided more consistency in test selection for this sample over time, but there would still have been a significant number of individuals who would not have achieved standard scores on this measure at any time point.

### Conclusions

The current study showed declines in nonverbal cognitive scores in a rare longitudinal sample of individuals with early diagnoses of ASD at all levels of cognitive ability. Our findings underscore that an understanding of cognitive ability trajectory in ASD is colored by methodological choices necessitated by characteristics of the population. Specifically, the core and associated deficits of ASD preclude the use of traditional age-appropriate IQ tests in many individuals with the disorder. Keeping such limitations in mind, we found that by age 3, NVIQ was generally

stable for the majority of individuals with ASD with IQs in the normal range. However, individuals exhibiting cognitive impairment by this early age were likely to decline further. It will be important for future studies to attempt to disentangle the influence of ASD symptoms and other potentially significant factors (e.g., verbal IQ, language abilities) that predict individual trajectories of cognitive development.

**Acknowledgments** This work was supported by Grants from the National Institute of Mental Health (MH081873), the National Institute of Child Health and Human Development (U 19 HD 035482), and Autism Speaks (dated January 11, 2008) to Catherine Lord. The funding sources played no role in the writing of the manuscript or the decision to submit it for publication, including study design, recruitment of the sample, or the collection, analysis and interpretation of the data. The authors were not paid by a pharmaceutical company to write this article. The authors had full access to all of the data in the study as well as the final responsibility for the decision to submit for publication. The views expressed in this paper do not necessarily represent the views of the NIMH, NIH, HHS, or the United States Government. The authors are grateful to all of the participating families in this study, as well as Pamela DiLavore, Cory Shulman, and Susan Risi for their critical roles in data collection. We thank Shanping Qiu for technical assistance with the data. We also gratefully acknowledge Catherine Lord for her input on the manuscript.

**Conflict of interest** The authors have no conflict of interest to report.

## References

- Aiken, L. R. (1996). *Assessment of intellectual functioning* (2nd ed.). New York: Springer.
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Publishing.
- Bishop, S. L., Guthrie, W., Coffing, M., & Lord, C. (2011). Convergent validity of the Mullen Scales of Early Learning and the differential ability scales in children with autism spectrum disorders. *American Journal on Intellectual and Developmental Disabilities, 116*(5), 331–343. doi:10.1352/1944-7558-116.5.331.
- Burgaleta, M., Johnson, W., Waber, D. P., Colom, R., & Karama, S. (2014). Cognitive ability changes and dynamics of cortical thickness development in healthy children and adolescents. *Neuroimage, 84*, 810–819. doi:10.1016/j.neuroimage.2013.09.038.
- CDC. (2012). Prevalence of autism spectrum disorders—Autism and developmental disabilities monitoring network, 14 Sites, United States, 2008 *MMWR CDC Surveillance Summaries* (Vol. 61, pp. 1–19).
- Chadwick, O., Cuddy, M., Kusel, Y., & Taylor, E. (2005). Handicaps and the development of skills between childhood and early adolescence in young people with severe intellectual disabilities. *Journal of Intellectual Disability Research, 49*(Pt 12), 877–888. doi:10.1111/j.1365-2788.2005.00716.x.
- Coplan, J., & Jawad, A. F. (2005). Modeling clinical outcome of children with autistic spectrum disorders. *Pediatrics, 116*(1), 117–122. doi:10.1542/peds.2004-1118.
- DiLavore, P. C., Lord, C., & Rutter, M. (1995). The pre-linguistic autism diagnostic observation schedule. *Journal of Autism and Developmental Disorders, 25*(4), 355–379.
- Duncan, A. W., & Bishop, S. L. (2013). Understanding the gap between cognitive abilities and daily living skills in adolescents with autism spectrum disorders with average intelligence. *Autism*. doi:10.1177/1362361313510068.
- Dykens, E. M., & Lense, M. (2012). *Intellectual disabilities and autism spectrum disorder: A cautionary note* (pp. 263–284). New York: Oxford.
- Eaves, L. C., & Ho, H. H. (2008). Young adult outcome of autism spectrum disorders. *Journal of Autism and Developmental Disorders, 38*(4), 739–747. doi:10.1007/s10803-007-0441-x.
- Eldevik, S., Hastings, R. P., Hughes, J. C., Jahr, E., Eikeseth, S., & Cross, S. (2010). Using participant data to extend the evidence base for intensive behavioral intervention for children with autism. *American journal on intellectual and developmental disabilities, 115*(5), 381–405. doi:10.1352/1944-7558-115.5.381.
- Elliot, C. (1990). *Manual for the differential abilities scales*. San Antonio, TX: Psychological Corporation.
- Elliott, C. D. (2007). *Differential ability scales* (2nd ed.). San Antonio, TX: The Psychological Corporation.
- Gillberg, C., & Steffenburg, S. (1987). Outcome and prognostic factors in infantile autism and similar conditions: A population-based study of 46 cases followed through puberty. *Journal of Autism and Developmental Disorders, 17*(2), 273–287.
- Gillespie-Lynch, K., Sepeta, L., Wang, Y., Marshall, S., Gomez, L., Sigman, M., et al. (2012). Early childhood predictors of the social competence of adults with autism. *Journal of Autism and Developmental Disorders, 42*(2), 161–174. doi:10.1007/s10803-011-1222-0.
- Girirajan, S., Dennis, M. Y., Baker, C., Malig, M., Coe, B. P., Campbell, C. D., et al. (2013). Refinement and discovery of new hotspots of copy-number variation associated with autism spectrum disorder. *American Journal of Human Genetics, 92*(2), 221–237. doi:10.1016/j.ajhg.2012.12.016.
- Greenspan, S., & Woods, G. W. (2014). Intellectual disability as a disorder of reasoning and judgement: The gradual move away from intelligence quotient-ceilings. *Current Opinion in Psychiatry, 27*(2), 110–116. doi:10.1097/ycp.0000000000000037.
- Hessl, D., Nguyen, D. V., Green, C., Chavez, A., Tassone, F., Hagerman, R. J., et al. (2009). A solution to limitations of cognitive testing in children with intellectual disabilities: the case of fragile X syndrome. *Journal of Neurodevelopmental Disorders, 1*(1), 33–45. doi:10.1007/s11689-008-9001-8.
- Howlin, P., Goode, S., Hutton, J., & Rutter, M. (2004). Adult outcome for children with autism. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 45*(2), 212–229.
- Howlin, P., Savage, S., Moss, P., Tempier, A., & Rutter, M. (2014). Cognitive and language skills in adults with autism: A 40-year follow-up. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 55*(1), 49–58. doi:10.1111/jcpp.12115.
- 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 and Allied Disciplines, 43*(6), 807–821.
- Kraijer, D. (2000). Review of adaptive behavior studies in mentally retarded persons with autism/pervasive developmental disorder. *Journal of Autism and Developmental Disorders, 30*(1), 39–47.
- Le Couteur, A., Lord, C., & Rutter, M. (2003). *The autism diagnostic interview—Revised (ADI-R)*. CA: Los Angeles.
- Lord, C., Risi, S., DiLavore, P. S., Shulman, C., Thurm, A., & Pickles, A. (2006). Autism from 2 to 9 years of age. *Archives of General Psychiatry, 63*(6), 694–701.
- Lord, C., Risi, S., Lambrecht, L., Cook, E. H, Jr, Leventhal, B. L., DiLavore, P. C., et al. (2000). The Autism Diagnostic Observation Schedule—Generic: A standard measure of social and communication deficits associated with the spectrum of autism. *Journal of Autism and Developmental Disorders, 30*(3), 205–223.



- Magiati, I., Tay, X. W., & Howlin, P. (2014). Cognitive, language, social and behavioural outcomes in adults with autism spectrum disorders: A systematic review of longitudinal follow-up studies in adulthood. *Clinical Psychology Review, 34*(1), 73–86. doi:[10.1016/j.cpr.2013.11.002](https://doi.org/10.1016/j.cpr.2013.11.002).
- Mullen, E. M. (1995). *Mullen Scales of Early Learning*. Circle Pines, MN: American Guidance Service Inc.
- Munson, J., Dawson, G., Sterling, L., Beauchaine, T., Zhou, A., Elizabeth, K., et al. (2008). Evidence for latent classes of IQ in young children with autism spectrum disorder. *American Journal of Mental Retardation, 113*(6), 439–452.
- Roid, G. H. (2003). *Stanford-Binet Intelligence Scales* (5th ed.). Itasca, IL: Riverside Publishing.
- Schalock, R. L., Verdugo, M. A., & Gomez, L. E. (2011). Evidence-based practices in the field of intellectual and developmental disabilities: an international consensus approach. *Evaluation and Program Planning, 34*(3), 273–282. doi:[10.1016/j.evalprogplan.2010.10.004](https://doi.org/10.1016/j.evalprogplan.2010.10.004).
- Sigman, M., & McGovern, C. (2005). Improvement in cognitive and language skills from preschool to adolescence in autism. *Journal of Autism and Developmental Disorders, 35*(1), 15–23.
- Sparrow, S. S., Cicchetti, D. V., & Balla, D. A. (2005). *Vineland adaptive behavior scales* (2nd ed.). Circle Pines, MN: AGS Publishing.
- Turner, L. M., Stone, W. L., Pozdol, S. L., & Coonrod, E. E. (2006). Follow-up of children with autism spectrum disorders from age 2 to age 9. *Autism, 10*(3), 243–265.
- Wechsler, D. (1999). *Wechsler abbreviated scale of intelligence*. San Antonio, TX: Psychcorp.
- Wechsler, D., & Naglieri, J. A. (2006). *Wechsler nonverbal scale of ability: Technical and interpretive manual*. Bloomington, MN: Pearson.