Language impairment in autistic adolescents and young adults: Variability by definition

Teresa Girolamo[1], Lindsay Butler[2], Julia Parish-Morris[3,4]

[1]San Diego State University, [2]University of Connecticut, [3]University of Pennsylvania School of Medicine, [4]Children's Hospital of Philadelphia

Key words: language impairment, autism, adolescents, adults, language assessment

Author Note

CRediT statement: TG: conceptualization, methodology, formal analysis, investigation, resources, data curation, writing – original draft, project administration, funding acquisition. LB: formal analysis, data curation, writing – review & editing, visualization. JPM: interpretation, writing – review and editing

Author Information: Teresa Girolamo, School of Speech, Language, and Hearing Sciences, San Diego, CA, San Diego State University. Lindsay Butler, Department of Speech, Language, and Hearing Sciences, Storrs, CT, University of Connecticut.

Funding: TG was supported by an ASHFoundation New Investigators Research Grant (PI: Girolamo). Research reported in this publication was supported by the National Institute on Deafness and other Communication Disorders of the National Institutes of Health under award number R21DC021769-01A. This award supported authors Girolamo and Butler to conduct this research. Research reported in this publication was supported by the National Institute on Deafness and other Communication Disorders of the National Institutes of Health under award number R01DC018289 (PI: Parish-Morris). This award supported co-author Parish-Morris to conduct this research. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Corresponding Author: Teresa Girolamo, 5201 Campanile Mall, San Diego State University, San Diego, CA, 07103. Phone: 619-594-4024. Email: tgirolamo@sdsu.edu.

Conflict of Interest: The authors have no financial or nonfinancial disclosures.

Data Availability Statement: Data are unavailable due to ethical restrictions. Information on data structure and analysis can be shared upon reasonable request to the corresponding author.

Abstract

Purpose: Though co-occurring structural language impairment (LI) in autism is common and predicts long-term outcomes, little is known about LI in autism beyond childhood. One challenge to closing this gap is that there is no consensus definition of LI. This study focuses on LI in autistic adolescents and young adults, asking to what extent clinical classification differs by definition and examining performance across language measures, nonverbal intelligence (NVIQ), and autism traits.

Method: Participants (*N* = 75; ages 13-30) varying in levels of autism traits completed norm-referenced measures of overall expressive language, overall receptive language, receptive vocabulary, expressive vocabulary, nonword repetition, and NVIQ. Scores were compared to epidemiological definitions for LI varying in criteria and cutoffs from -1 *SD* to -1.5 *SD*. Data were analyzed using descriptives and clustering.

Results: More stringent definitions yielded a greater proportion of participants meeting LI criteria, and more stringent cutoffs for LI yielded greater overall consistency in clinical classification across individual language measures, but there was no one-to-one ratio between cutoff and clinical classification. Clustering indicated three profiles differentiated on the basis of language and nonverbal cognitive skills, but each cluster was heterogeneous. Individual performance also varied across language measures.

Discussion: Findings support multi-domain approaches to characterizing language skills in autistic adolescents and adults, including those with LI. Future work is needed to understand language skills in autism beyond childhood and how to develop effective assessment practices.

Language impairment in autistic adolescents and young adults: Variability by definition

There is growing evidence of heterogeneity in language across the autism spectrum (Bal et al., 2016; Butler et al., 2023; Pizzano et al., 2024), but little is known about language in autism beyond childhood (Howlin & Taylor, 2015). Language skills in autism are tied to long-term academic, mental and physical health, occupational, and social outcomes (Brignell et al., 2018; Magiati et al., 2014). In the United States (U.S.), a key stage is the transition to adulthood (Roux et al., 2024), as autistic youth lose eligibility for child- and school-based services ("Individuals with Disabilities Education Improvement Act of 2004," 2018). Many lack access to appropriate services and supports during this time (Eilenberg et al., 2019; Weir et al., 2022), which is particularly relevant for the more than 50% of all autistic individuals with structural language impairment (LI), or challenges in morphology, phonology, and syntax, and the 63% with overall LI, or general language delays (Baird et al., 2006; Boucher, 2012; Levy et al., 2010). The first step to reducing these barriers is improving language assessment for older individuals, which provides evidence of meeting eligibility criteria for services (Anderson & Butt, 2018).

For transition-aged autistic individuals, characterizing language skills requires attention and care. Prior work has focused on broad language subgroups, such as LI or “minimally verbal” (Girolamo et al., 2023; Schaeffer et al., 2023). While such labels can facilitate access to services, they do not provide precise information about language skills (Rosen et al., 2021). Under current diagnostic criteria, autism can co-occur with LI and intellectual disability (ID) (American Psychiatric Association, 2013); language in autism can also dissociate from cognitive skills (Bal et al., 2016; World Health Organization [WHO], 2022). Detecting these patterns requires assessing multiple dimensions of individual skills within the domains of language and cognition (Kover & Abbeduto, 2023). To better understand the relationship between clinical classification

and linguistic variability, this report examines epidemiological definitions for LI and language assessment performance across domains in speaking autistic adolescents and adults.

Structural LI in Autistic Adolescents and Adults

There are known challenges in characterizing LI in autism beyond childhood. While there are validated clinical language assessments for youth (Girolamo et al., 2022; Nitido & Plante, 2020), there are few measures for adults over age 21 (Manenti et al., 2023). Understanding outcomes on norm-referenced assessments is important, as they are often part of service eligibility criteria within the U.S. (Burke et al., 2024; Selin et al., 2022). Unfortunately, existing adult language assessments are inadequate. Some were normed decades ago on non-representative samples of adults with a college education (Hammill et al., 2007) or assess a single language domain that underestimates LI, such as vocabulary (Calder et al., 2023; Schaeffer et al., 2023). There are several other limitations.

First, studies vary in how they characterize language skills in adults, which hinders generalizability and validity. Some studies described language in autistic adults using norm-referenced tests designed for other populations. For example, receptive and expressive vocabulary tests for younger individuals, the British Picture Vocabulary Scale Picture (Dunn et al., 1982) and Expressive One-Word Picture Vocabulary Test (Gardner, 1982), yielded age equivalent scores outside the possible range: 5-11% below the basal and 14-20% above the ceiling (N = 60; ages 29-64; M_{PIQ} = 90-93) (Howlin et al., 2014), nearly 50% below the basal (N = 19; ages 21-26; M_{PIQ} = 73) (Mawhood et al., 2000), and 16% below the basal and above ceiling (N = 68; ages 21-48; M_{PIQ} = 75) (Howlin et al., 2004). Further, using overall receptive-expressive language tests for adults with aphasia, or the Western Aphasia Battery (Kertesz, 1982) and Right Hemisphere Language Battery (Bryan, 1989), to assess autistic adults without brain damage (N =

17; ages 18-67; M_{NVIQ} = 91) leaves unclear how to interpret findings (Lewis et al., 2008).

Other studies defined LI in adults using age 21 norms for older individuals, with varying criteria and cutoffs from -1 *SD* to < -2 *SD*. In nonautistic adults, LI definitions have included: (a) -2 *SD* on the Clinical Evaluation Language Fundamentals (CELF)-4 (Semel et al., 2006) overall expressive-receptive language (*N* = 83; age 24; M_{NVIQ}: 99), as -2 *SD* indicates communicative challenges (Botting, 2020); (b) -1.5 *SD* on CELF-5 UK Core Language (Semel et al., 2017) (*N* = 44; ages 16-24; IQ not assessed) (Clegg et al., 2021), per a Delphi consensus study (Bishop et al., 2017) and an epidemiological study of LI in nonautistic youth (Norbury et al., 2016); and (c) -1 *SD* on receptive vocabulary or an expressive-receptive language composite, or -2 *SD* on ≥ 2 subtests of an omnibus language measure (*N* = 13; ages 18-25; PIQ ≥ 80) (Poll et al., 2010), per local norms developed for another sample (Johnson et al., 1999b). In autistic adults (*N* = 13; ages 17-23; M_{NVIQ}: 80), LI has been defined as -1.25 *SD* on ≥ 2 measures: CELF-5 Receptive Language, CELF-5 Expressive Language (Wiig et al., 2013a), Peabody Picture Vocabulary Test (PPVT)-5 (Dunn, 2019), Expressive Vocabulary Test (EVT)-3 (Williams, 2019), and Syllable Repetition Task (SRT) percent accuracy (Shriberg et al., 2009b) (Girolamo & Rice, 2022), per epidemiological study of LI in nonautistic youth (Tomblin et al., 1996). Beyond differences in sampling, as epidemiological definitions (b), (c), and (d) were developed for samples in the United Kingdom, Canada, and United States, there is no consensus on how to define LI in adults.

Relatedly, independent validation studies reveal no single definition for LI in adults. In the epidemiological Ottawa study, differences from test norming samples motivated the development of local norms to define LI (*N* = 114; ages 18-20; M_{PIQ} = 92-105), inclusive of autism (Johnson et al., 1999b): (a) -1 *SD* of the PPVT-Revised (Dunn & Dunn, 1981) or Test of Adolescent/Adult Language-3 (TOAL-3) Spoken Language Quotient (Hammill et al., 1994), or

(b) -2 *SD* on ≥2 TOAL-3 subtests (Johnson et al., 1999a). The -1 *SD* cutoff yielded higher LI estimates (12.6%) than expert clinical judgment (8.3%) (Johnson et al., 1999a). In a separate sample, a combination of assessments yielded the highest classification accuracy of LI in nonautistic adults with a history of developmental LI (N = 96; ages 18-49; M_{NVIQ} = 97-103): an 11-word spelling measure for phonology (Benton & Hamsher, 1978), CELF-4 Word Definitions for semantics (Semel E et al., 2003), and the Modified Token Test (Morice & McNicol, 1985) for syntax (Fidler et al., 2011, 2013). Single measures of narration, nonword repetition, sentence production, and grammatical judgment were inconsistent in identifying LI (Fidler et al., 2011). While narration can confound structural language skills with autism traits (Baixauli et al., 2016)—which prior work has found to be distinct from LI (Loucas et al., 2008; Manenti et al., 2024; Silleresi et al., 2020)—findings support the clinical utility of multi-domain, multi-method assessment.

A third limitation is many studies on LI in autism use an IQ cutoff (Girolamo et al., 2023), which does not reflect current autism diagnostic criteria (WHO, 2022) or the 39.6% of autistic individuals with intellectual disability (Russell et al., 2019; Shaw et al., 2025). A recent study found structural language and NVIQ dissociated in 25% of 38 autistic adults (ages 18-56 years), with three heterogeneous groups: 1) typical language and average or above NVIQ (n = 23), though some had moderate LI on sentence repetition (n = 2) or severe LI on sentence repetition and moderate LI on nonword repetition (n = 1); 2) low NVIQ with LI (n = 6) or with typical language (n = 3); and 3) LI with low NVIQ (n = 6) (Manenti et al., 2024). To note, NVIQ was classified as low (< 80), average (80-109), or high (≥110) based on the Wechsler Adult Intelligence Scales-IV (Wechsler, 2008) Block Design and Matrix Reasoning Perceptual Reasoning Index (n = 38) (Manenti et al., 2024). LI was classified as moderate at -1.25 *SD* and -

1.65 *SD* the mean percent accuracy of 50 nonautistic adults without prior psychiatric diagnoses or hearing loss (ages 18-73) on non-normed measures of nonword repetition (Ferré et al., 2022) and sentence repetition (Tuller et al., 2021) (Manenti et al., 2024). Altogether, heterogeneity in LI definitions and outcomes motivate a dimensional approach that can improve precision in language assessment.

The Current Study

In the absence of robust evidence, examining patterns in assessment helps build the evidence base (Oetting & McDonald, 2002). This exploratory, observational study uses a comprehensive protocol of direct behavioral language assessments, which enhances confidence in findings (Arunachalam et al., 2022). Further, this study examined epidemiological LI definitions using normed measures (Johnson et al., 1999a; Norbury et al., 2016; Tomblin, 1996), facilitating comparisons. Last, given current autism diagnostic criteria (WHO, 2022) and systematic exclusion in autism research (Russell et al., 2019), this study had no NVIQ cutoffs. Research questions were:

1) To what extent does clinical classification of LI differ by definition: (a) -1 *SD* on omnibus language or receptive vocabulary measures, or -2 *SD* on ≥2 subtests of an omnibus language measure (Johnson et al., 1999a); (b) -1.25 *SD* on ≥2 measures of overall receptive language, overall expressive language, receptive vocabulary, expressive vocabulary, and nonword repetition, or on an omnibus language measure (Tomblin et al., 1996); and c) -1.5 *SD* on ≥2 measures from (b) (Norbury et al., 2016), and by individual language measure?
2) How does performance on language measures and NVIQ compare?

Given higher concordance between clinical judgment and lower language scores (Johnson et al., 1999a), it was expected that more stringent definitions would yield lower LI

estimates and be lower for definition (c) than (a) or (b). As vocabulary may be a relative linguistic strength in autism (Arunachalam & Luyster, 2016), it was also expected fewer participants would meet LI cutoffs on receptive and expressive vocabulary than on other measures (Schaeffer et al., 2023). Last, per current autism diagnostic criteria (WHO, 2022) and prior work (Manenti et al., 2024), it was expected that language and NVIQ would dissociate in some participants.

Methods

Participants

Selection criteria were: (a) meeting DSM-5 autism diagnostic criteria (APA, 2013), per presence of a professional autism diagnosis and confirmation using the Social Responsiveness Scale-2 (SRS-2) (Constantino & Grubler, 2012) plus expert clinical judgment; (b) ages 13 to 30, coinciding with the onset of transition planning in some states within the U.S. and 10 years post-eligibility for special education services ("Every Student Succeeds Act," 2015; "Individuals with Disabilities Education Act," 2004); (c) English as a first language, as study activities were in English; (d) adequate hearing and vision thresholds for hearing human speech and seeing computer screen images, as study activities required responding to audiovisual stimuli, and; (e) use of primarily spoken language to communicate, as activities required oral responses.

Procedures

This study received institutional board approval and recruited participants using a multi-step process. The team shared study flyers with organizations providing services to autistic individuals, provided personalized consultation about the study to potential participants and caregivers, and obtained informed consent and assent prior to data collection. Participants provided informed consent if they were their own legal guardian; otherwise, caregivers provided

consent and participants provided assent. Recruitment and data collection took place from 2022 to 2024 remotely on HIPAA-compliant Zoom. Participants completed assessments using the digital version of assessments and test developer guidance (Pearson, 2025).

Measures

Demographics

Participants or caregivers provided information on participant race and ethnicity following U.S. Census categories: American Indian, Native American or Alaska Native, Asian, Black or African American, Pacific Islander, white, other, or more than one race, and Hispanic or Latino (yes/no) (Office of Management and Budget, 1997). Respondents could also select if they did not know race or ethnicity, as well as write in options. Similarly, respondents provided information on participant sex at birth and gender, with the ability to write in options.

Autism Traits

Autism traits were assessed using SRS-2 caregiver or self-report forms for students and adults (α: .95-.97) (Constantino, 2012). Respondents indicate the frequency of 65 items on a four-point scale, yielding an overall *t*-score of subclinical (59), mild (60-65), moderate (66-75), or high levels (≥76) of autism traits. Per best practices (Bishop & Lord, 2023), autism status was determined by triangulating formal autism diagnosis, SRS-2 *t*-scores, and expert judgment by a research reliable practitioner on the Autism Diagnostic Observation Schedule-2 (Lord et al., 2012). As SRS-2 means of caregiver student forms (n = 37), caregiver adult forms (n = 15), and adult self-report forms (n = 21) non-significantly differed (p = .061), forms were combined.

Language Skills

Participants completed untimed assessments of nonword repetition, omnibus language, and receptive-expressive vocabulary. Nonword repetition was assessed by SRT percent accuracy

(in 16-year-olds, the oldest norming group, *M*: 93.2, *SD*: 3.0; *r*=.69-.89) (Shriberg et al., 2009b; Shriberg & Mabie, 2017). Participants repeat 18 nonwords of two to four syllables, which use early-acquired sounds (/b/, /d/, /m/, /n/, and /a/) to remove confounds due to misarticulation (Shriberg et al., 2009a). Omnibus language was assessed by six subtests of the CELF-5 (*M*: 10, *SD*: 3), valid for ages 3 to 21 (Wiig et al., 2013a). Expressive language subtests were: (a) Formulated Sentences (*r* = .86), which requires producing sentences of increasing semantic and syntactic complexity in response to a visual and auditory single-word stimulus; (b) Recalling Sentences (*r* = .94), which requires repeating sentences of increasing semantic, syntactic, and morphological complexity verbatim; and (c) Semantic Relationships (*r* = .89), which requires completing a sentence by selecting two of four semantically related words or phrases. Receptive language subtests were: (a) Word Classes (*r*: .90), which requires selecting two of three or four words based on associated classes; (b) Understanding Spoken Paragraphs (*r* = .85), which requires listening to stories of increasing complexity and verbally answering questions; and (c) Sentence Assembly (*r* = .93), which requires building sentences using audiovisual stimuli. Subtest scaled scores (*M* = 10, *SD* = 3) yield composite standard scores for expressive language, receptive language, and overall expressive-receptive language (*r* = .95-.96; *M* = 100, *SD* = 15). This study used age 21 norms for participants over age 21 (*n* = 24), per prior work (Botting, 2020; Clegg et al., 2021; Fidler et al., 2011). Receptive vocabulary was assessed by the PPVT-5 (Dunn, 2019) and expressive vocabulary by the EVT-3 (Williams, 2019), which are co-normed measures valid for ages 2.5 to >90 years (*r* = .97; *M* = 100, *SD* = 15). On the PPVT-5, participants select one of four images in response to an auditory stimulus word. On the EVT-3, participants provide one word in response to an auditory question and visual stimulus.

NVIQ

NVIQ was assessed using the digital long form of the Raven's Progressive Matrices-2nd Ed. (r = .88-.91; M = 100, SD = 15), which is untimed and valid for ages 4 to 90 years (Raven et al., 2018). Participants view a stimulus picture and select an image from a field of five to complete the picture. The Raven's does not use language, reducing a clinically relevant confound for autistic individuals and especially for those with LI (Grondhuis et al., 2018).

Data Processing and Analysis

A trained research assistant independently checked measures except for NVIQ and SRS-2, which were autoscored on their respective platforms. Next, data were examined for missingness, with plans to use means matching with a single imputation if < 5% of data was missing or multiple imputation if >5% of data was missing (Little & Rubin, 2019). For the first research question, participants were classified as meeting each LI definition (yes/no): (a) -1 *SD* on CELF-5 core language or PPVT-5, or < -2 *SD* on ≥2 CELF-5 core language subtests (Johnson et al., 1999a); (b) -1.25 *SD* on ≥2 measures: CELF-5 Receptive Language Index, CELF-5 Expressive Language Index, PPVT-5, EVT-3, and SRT accuracy (Tomblin et al., 1996); and (c) -1.5 *SD* on ≥2 measures from (b) (Norbury et al., 2016). Participants were also classified as meeting each LI cutoff of -1 *SD*, -1.25 *SD*, or -1.5 *SD* on each language measure (yes/no): PPVT-5, EVT-3, SRT accuracy, and CELF-5 subtests (Formulated Sentences, Recalling Sentences, Semantic Relationships, Word Classes, Understanding Spoken Paragraphs, Sentence Assembly). For the second research question, autism trait, language, and NVIQ scores were *z*-scored. Participants were classified by NVIQ band per test developer and clinical guidelines indicating that an IQ of 70 to 75 is clinically significant (American Association on Intellectual and Developmental Disabilities, 2024; Raven et al., 2018): low (≤75), borderline (76-89), average (90-109), and high (≥110).

All analyses were conducted using SPSS 31.0 (IBM Corp., 2025) using an α level of .05. As data violated assumptions of normality, analyses used nonparametric measures. To address the first research question, Cochran's Q (Cochran, 1950) tested to what extent LI classification differed by definition. To test if LI classification differed by language measure, pairwise comparisons were conducted using Dunn's (1964) procedure, with a Bonferroni (1936) correction for multiple comparisons. Prior to analysis, data were checked for random sampling and sampling adequacy.

Next, to examine patterns between language and NVIQ, given the sample size (Dalmaijer et al., 2022), data were analyzed using principal component analysis (PCA) and clustering, an unsupervised machine learning method appropriate for exploratory analysis (Husson et al., 2010). A PCA was run on SRS-2 social communication and interaction, SRS-2 restricted and repetitive behavior, CELF-5 language subtest, PPVT-5, EVT-3, and SRT *z*-scores (Abdi & Williams, 2010). Prior to analysis, variables were examined for multicollinearity using a cutoff of variance inflation factor ≥ 10, and the PCA was checked for linearity, sampling adequacy, and outliers. Then, agglomerative hierarchical clustering, where each data point starts as its own cluster and is iteratively merged with the most similar clusters until there is one cluster, using Ward's method (1963) and a squared Euclidean distance interval was run on language and NVIQ *z*-scores (Bouguettaya et al., 2015; Day & Edelsbrunner, 1984). Inspection of the dendrogram and agglomerative schedule were used to determine a number of solutions and initial centers for *k*-means clustering (Lu et al., 2008), which iteratively groups data points into *k* number of clusters to minimize distance between data points and cluster centroids (Hartigan & Wong, 1979). *K*-means clustering was run on language and NVIQ *z*-scores, with a Euclidian distance interval, 999 iterations, and no running means. Internal cluster validity was evaluated using the SPSS

package KO_MACROS (Orlov, 2024). External validity was evaluated using discriminant

Table 1

Sample Characteristics (N = 75)

function analysis, Cohen's κ (1960) of *k*-means and hierarchical clustering outcomes, and descriptive statistics (Dudoit & Fridlyand, 2002). Effects were interpreted as: small at .25, medium at .55, and large at .95 (Gaeta & Brydges, 2020).

Results

Sample Characteristics

Recruitment resulted in a sample of 75 participants (ages 13-30 years); see Tables 1 and 2. Data were missing from three participants: SRS-2 scores (*n*=2; forms missing) and language and NVIQ scores (*n*=1; missing). Given the focus of this study, the participant with missing language and NVIQ scores was excluded from analysis. Given the small amount of missing data (Little & Rubin, 2019), autism trait scores were replaced using means matching with one imputation. The sample varied in racial and ethnic background, with a higher female-to-male ratio for sex at birth and gender than the overall estimate of 1:3.4 (Shaw et al., 2025), whether considering the overall sample (1:2-2.1), participants 21 years of younger (1:2.2-2.3), and those over age 21 years (1:1.6). Overall, the sample had a moderate level of autism traits.

Variable	Full Sample (N = 75)		≤ 21 years (n = 52)		> 21 years (n = 23)	
	n	%	*n*	%	*n*	%
Age	75	19.70 (4.67), 13.27-30.43	52	17.11 (2.61), 13.27-21.75	23	25.55 (2.41), 22.20-30.43
Race						
American Indian, Native American, Alaska Native	12	16.00	6	11.54	6	26.07
Asian	19	25.33	16	30.77	3	13.04
Black	39	52.00	26	50.00	13	56.52
multiracial	26	34.67	18	34.62	8	34.78
Pacific Islander	4	5.33	3	5.77	1	4.35
white	21	28.00	14	26.92	7	30.43
other	9	12.00	7	13.46	2	8.70
Hispanic/Latine: Yes	15	20.00	11	21.15	4	17.39
Sex at Birth: Female-to-Male	75	32 : 68	52	28.85 : 71.15	23	39.13 : 60.87
Gender: Female-to-Male	75	33.33 : 66.67	52	30.77 : 69.23	23	39.13 : 60.87

Note. Age presented as *M* (*SD*), and range. All other variables are presented as percentages. Multiracial is non-exclusive with other categories. Other: Latine (n = 7), Middle Eastern (n = 1), and Puerto Rican (n = 1).

Table 2

Sample Clinical Assessment Scores (N = 75)

Variable	Full Sample (N = 75)			≤ 21 years (n = 51)			> 21 years (n = 23)		
	n	M (SD)	range	n	M (SD)	range	n	M (SD)	range
	7	96.26	47-		87.90	47-	23	95.48	66-126
Raven's NVIQ standard score	4	(16.46)	135	51	(16.99)	135		(15.35)	
SRS-2 social communication impairment	7	71.56	50-90	51	70.80	50-90	22	73.36	57-90
t-score	3	(11.19)			(11.94)			(9.28)	
SRS-2 restricted & repetitive behaviors t-	7	72.07	43-90	51	69.86	43-90	22	77.18	55-90
score	3	(12.85)			(13.29)			(10.35)	
	7	76.85	45-	51	75 (18.87)	45-	23	79.35	45-110
CELF-5 Expressive Language Index	4	(16.29)	116			116		(20.31)	
	7	80.86	45-	51	77.76	45-	23	87.78	48-118
CELF-5 Receptive Language Index	4	(21.83)	118		(19.66)	114		(22.73)	
Peabody Picture Vocabulary Test-5	7	90.59	48-	51	89.92	48-	23	92.09	56-123
standard score	4	(21.41)	129		(21.77)	129		(20.07)	
Expressive Vocabulary Test-3 standard	7	92.54	48-	51	91.67	49-	23	94.48	57-128
score	4	(18.97)	128		(18.49)	122		(19.73)	
	7		1-14	51	6.27 (3.85)	1-14	23	6.91 (3.58)	1-13
CELF-5 Recalling Sentences scaled score	4	6.47 (3.74)							
	7	88.54	48-	51	88.22	48-	23	98.26	66-100
Syllable Repetition Task percent accuracy	4	(18.54)	100		(18.97)	100		(18.74)	

Note. Peabody Picture Vocabulary Test-5th Ed. (Dunn, 2018). Expressive Vocabulary Test-3rd Ed. (Williams, 2018). CELF-5 – Clinical Evaluation of Language Fundamentals-5th Ed. (Wiig et al., 2013). Syllable Repetition Task (Shriberg et al., 2009). Two participants had missing SRS-2 scores. One participant had missing language and NVIQ scores.

LI Classification by Definitions and Measure

Overall, there was a significant decrease in clinical classification of LI from less to more stringent definitions, $\chi^2(2)=19.54$, $p < .0001$: (a) 70.27% per -1 *SD* on overall language or receptive vocabulary, or -2 *SD* on ≥2 overall language subtests (Johnson et al., 1999a); (b) 62.16% per -1.25 *SD* on ≥2 measures of overall receptive language, overall expressive language, receptive vocabulary, expressive vocabulary, and nonword repetition (Tomblin et al., 1996); and (c) 52.70% per -1.5 *SD* on ≥2 measures of overall receptive language, overall expressive language, receptive vocabulary, expressive vocabulary, and nonword repetition (Norbury et al., 2016). Yet only the decrease from (a) to (c) was significant (17.57%), $p < .0001$. Thus, there was no one-to-one ratio between stringency of LI definition and classification.

More stringent cutoffs generally yielded greater consistency in LI classification across language measures, as indicated by fewer significant pairwise comparisons: -1 *SD* (n = 18), -1.25 *SD* (n = 13), and -1.5 *SD* (n = 9); see Supplementary Table 1. Yet outcomes varied by measure; see Figures 1 and 2. Understanding Spoken Paragraphs had the most significant pairwise comparisons at -1 *SD* and -1.25 *SD*; more participants met these LI cutoffs than on expressive vocabulary (Δ: 33.78-40.54%), receptive vocabulary (Δ: 25.68-35.13%), Sentence Assembly (Δ: 18.92-29.73%), or Word Classes (Δ: 21.62-32.43%). Also, more participants met LI cutoffs on overall receptive language than expressive vocabulary at -1.25 *SD* and -1.5 *SD* (Δ: 22.97%) or receptive vocabulary at -1.5 *SD* (Δ: 20.27%). Expressive vocabulary had the most significant pairwise comparisons of all language measures at -1.5 *SD*, with fewer participants meeting this LI cutoff relative to overall expressive language (Δ: 22.97%), overall receptive language (Δ: 22.97%), or sentence production (Δ: 29.73%). More broadly, fewer participants met any LI cutoff on expressive vocabulary than on overall expressive language (Δ: 22.97-32.44%) and sentence repetition (Δ: 20.27-21.62%). In all, no one measure reliably classified participants.

Language Performance in NVIQ Patterns

To examine patterns between language and NVIQ, a PCA was run on autism traits, language, and NVIQ *z*-scores. All variables had at least one $r > 0.3$, the overall Kaiser-Meyer-Olkin measure was 0.89, or “meritorious” (Kaiser, 1974), and Bartlett’s (1951) test of sphericity was significant ($p < .0001$), indicating data was likely factorizable. There were two components with eigenvalues >1, which explained 58.48% and 14.81% of the total variance. Inspection of the scree plot and the interpretability criterion indicated two components should be retained (Cattell, 1966); see Figure 3. The two-component solution explained 73.28% of the total variance. Using a Varimax orthogonal rotation for interpretability, the rotated solution exhibited simple structure,

with strong loadings of language and NVIQ on Component 1 (50.32%) and autism traits on Component 2 (24.99%) ; see Supplementary Table 2. As autism traits weakly correlated with Component 1 ($r < .16$), clustering was run on only language and NVIQ.

Agglomerative hierarchical clustering suggested there were three clusters and provided initial cluster centers; see Figure 4 and Supplementary Table 3. Next, *k*-means clustering was run on language and NVIQ *z*-scores. Internal clustering criteria indicated three clusters was optimal, which external validity indices also supported; see Supplementary Table 4. Discriminant function analysis accurately classified 100% of participants from observed data and 93.24% in cross-validation. Of 14 participants expected to be in Cluster 1, one was in Cluster 3; of 31 expected to be in Cluster 2, one was in Cluster 3, and; of 29 expected to be in Cluster 3, three were in Cluster 2. Clustering methods also had near-perfect agreement, $\kappa = .873$, 95% CI [.775, .971], $p < .0001$. Thus, a three-cluster solution was selected and interpreted.

Descriptively, clusters differed on the basis of mean language and NVIQ scores, as well as representation in terms of sex at birth; see Figure 5 and Table 3. Female-to-male ratios ranged from 1.58 in Cluster 2 to 2.63 in Cluster 3 and 3.67 in Cluster 1. Cluster 1 ($n = 14$) had language scores < -2 *SD* and NVIQ < -1.5 *SD*, Cluster 3 ($n = 29$) had language scores < -1 *SD* and NVIQ of approximately -1 *SD*, and Cluster 2 ($n = 31$) had language scores ≥ -1 *SD* and NVIQ at the mean. However, performance across clusters was heterogeneous; see Figure 6 and Supplementary Figure 1. For example, if considering clinical categories, there was overlap in language and NVIQ: (a) < -1.5 *SD* on ≥2 measures and low NVIQ (Cluster 1: $n = 6$, Cluster 3: $n = 5$), (b) < -1.5 *SD* (Cluster 1: $n = 5$, Cluster 2: $n = 1$, Cluster 3: $n = 4$) or < -1.25 *SD* (Cluster 3: $n = 1$, Cluster 2: $n = 1$) on ≥2 measures and borderline NVIQ, (c) < -1.5 *SD* on ≥2 measures and average NVIQ (Cluster 1: $n = 3$, Cluster 3: $n = 11$, Cluster 2 = 3), and (d) < -1.25 *SD* on ≥2

measures and high NVIQ (Cluster 3: *n* = 1, Cluster 2: *n* = 1). Individual performance also varied within each cluster.

Table 3

K-Means Three-Cluster Solution Cluster Means, Minimum, and Maximum Scores

Measure	Cluster 1 (*n*=14)	Cluster 2 (*n*=31)	Cluster 3 (*n*=29)
Age	20.03 (13.48-29.67)	20.26 (13.35-30.43)	19.09 (13.27-29.03)
Sex at birth: female-to-male	3:11	12:19	8:21
NVIQ standard score	76.36 (47-100)	100.55 (78-135)	85.97 (57-111)
Receptive vocabulary standard score	62.36 (40-89)	108.13 (86-129)	85.48 (57-117)
Expressive vocabulary standard score	67.64 (49-87)	109.68 (85-128)	86.24 (65-106)
Nonword repetition percent accuracy	75.23 (48-96)	94.52 (80-100)	90.36 (68-100)
Word Classes scaled score	2.79 (1-9)	11.97 (8-14)	6.03 (1-12)
Formulated Sentences scaled score	2.07 (1-7)	8.23 (3-14)	3.93 (1-11)
Recalling Sentences scaled score	2.43 (1-7)	9.52 (2-14)	5.16 (1-9)
Understanding Spoken Paragraphs scaled score	2.36 (1-7)	7.61 (4-11)	3.86 (1-9)
Sentence Assembly scaled score	2.43 (1-8)	9.55 (4-18)	5.72 (1-9)
Semantic Relationships scaled score	2.93 (1-8)	10.87 (6-16)	5.48 (1-9)

Note. All variables presented as means, minimum, and maximum. NVIQ, receptive vocabulary, and expressive vocabulary standard score *M* = 100, *SD* = 15. Nonword repetition percent accuracy: *M* = 93.2, *SD* = 3.0. Word Classes, Formulated Sentences, Recalling Sentences, Understanding Spoken Paragraphs, Sentence Assembly, and Semantic Relationships scaled score: *M* = 10, *SD* = 3. Participants over 21: *n* = 5 in Cluster 1, *n* = 13 in Cluster 2, and *n* = 5 in Cluster 3.

In Cluster 1, nearly all participants scored < -1.5 *SD* on all language measures (*n* = 9, or 64.29%) or all but two of nine measures (*n* = 3, or 21.43%). Exceptions to Cluster 1 included one participant with mostly typical scores, except for scores of < -1.5 *SD* (sentence production, nonword repetition) and of < -1 *SD* (sentence repetition, Understanding Spoken Paragraphs) on two measures each. The other Cluster 1 participant had mostly typical scores (receptive vocabulary, expressive vocabulary, Semantic Relationships) or scores of < -1 *SD* (sentence

repetition, Sentence Assembly), with remaining scores at < -1.5 *SD*. Similar to Cluster 1, while most Cluster 3 participants scored < -1.5 *SD* on ≥2 measures (*n* = 26, or 89.66%), there were exceptions. Three (10.34%) scored < -1.25 *SD* on two to three measures: < -1.25 *SD* on sentence repetition and < -1.5 *SD* on Understanding Spoken Paragraphs; < -1.25 *SD* on sentence assembly and < -1.5 *SD* on sentence repetition; and < -1 *SD* on nonword repetition, < -1.25 *SD* on sentence production, and -1.5 *SD* on Understanding Spoken Paragraphs. In contrast, over two-thirds of Cluster 2 participants had typical scores on all language measures (*n* = 11, or 35.48%) or all but one measure (*n* = 12, or 38.71%). In this latter group, scores were mostly at < -1.5 *SD* and commonly included sentence production (*n* = 4), nonword repetition (*n* = 3), and Understanding Spoken Paragraphs (*n* = 3). Other Cluster 2 participants scored mostly < -1.25 *SD* on two to three measures (*n* = 6, or 20.69%), including: sentence production (*n* = 4), nonword repetition (*n* = 3), and Understanding Spoken Paragraphs (*n* = 3). Overall, findings point to variability, regardless of broad group (i.e., cluster) characteristics.

Discussion

Analyzing language assessment performance in speaking autistic adolescents and young adults ranging in NVIQ and levels of autism traits revealed differences in clinical classification for LI by definition. Exploratory findings from clustering analyses also provide information to guide current interpretation and use of language assessment. Sample heterogeneity at individual and group levels has implications for understanding language beyond childhood and improving precision in language assessment to support service delivery.

LI Classification and Assessment Performance

Findings partly replicated prior work. Three LI definitions yielded clinical classification rates of 52.70% (Norbury et al., 2016), 62.16% (Tomblin et al., 1996), and 70.27% (Johnson et

al., 1999a). On one hand, 52.70% and 62.16% are comparable to rates of structural LI (>50%) and overall LI (63%) in epidemiological samples of autistic youth (ages 8-10 years) (Baird et al., 2006; Levy et al., 2010). However, they are higher than rates from a recent sample of 38 autistic adults, where 26.32% scored < -1.25 *SD* on two measures (Manenti et al., 2024). This discrepancy could be due to differences in sampling and design. First, this study had a younger sample (ages 13-30 versus 18-56 years) and used less stringent LI cutoffs (< -1 *SD* to < -1.5 *SD* versus < -1.25 *SD* and < -1.65 *SD*). Similarly, this study used different approaches to scoring: test norms for most measures and maximum age test norms for omnibus language and nonword repetition versus *z*-scores from a "control group" of 50 nonautistic adults (Manenti et al., 2024). LI rates might be lower in this sample if using more stringent cutoffs or different norms for the 24 participants over age 21. Second, this study used additional measures beyond nonword repetition and sentence repetition (Manenti et al., 2024), per LI definitions based on primarily nonautistic samples (Johnson et al., 1999b; Norbury et al., 2016; Tomblin et al., 1996). Inclusion of measures, such as overall language subtests, might reflect pragmatic differences (Creemers, 2015) and lead to higher LI classification rates (Baird et al., 2006). Third, this study used an NVIQ measure that did not rely on language for following directions, with differences in low NVIQ rates (16.2% versus 23.68%) and cutoffs (≤75 versus < 80) (Manenti et al., 2024). These differences in NVIQ measures and rates could have influenced LI classification.

Still, the LI classification rate of 70.27%, which uses a cutoff of -1 *SD* (Johnson et al., 1999a), is higher than estimates from younger autistic children (Baird et al., 2006; Levy et al., 2010). Beyond pragmatic differences, one possibility is that child-based criteria for LI do not map onto older individuals, amid lack of consensus about how to define LI. To this end, clinical judgment of LI status was more stringent than local norms for adolescents and young adults

(ages 18-20), especially for those scoring close to the -1 *SD* cutoff on receptive vocabulary or overall language; however, clinicians judged those scoring < -1 *SD* on both measures as having LI (Johnson et al., 1999b). LI criteria may also explain differences. Here, there was only a significant decrease in LI classification between the least and most stringent LI definitions in terms of cutoff (-1 versus -1.5 *SD*) and measures (receptive vocabulary or overall language versus a wider array of measures). In Fidler et al. (2011), individual language measures varied in their discriminatory ability for nonautistic adults with LI (ages 18-49; sensitivity: 35-75%, specificity: 53-87%), and did not replicate LI criteria from an epidemiological sample of nonautistic adolescents (ages 13-17) (Tomblin, 2008). Overall, findings do not support a singular definition for LI at this point in time.

Consistent with expectations (Schaeffer et al., 2023), single-domain expressive and receptive vocabulary measures yielded lower LI classification rates relative to other measures. However, only some differences were statistically significant. Notably, rates were not lower for receptive vocabulary than sentence repetition or nonword repetition, nor for expressive vocabulary than nonword repetition. Whether these statistically significant differences are clinically significant is unknown (Chatham et al., 2018). Determining this information requires a “gold standard,” which is elusive for older individuals (Fidler et al., 2011), let alone autistic young adults (Howlin & Taylor, 2015). While the exploratory nature of this study does not support determining profiles, clusters do provide information on patterns (Oetting & McDonald, 2002). As in prior work (Manenti et al., 2024; Silleresi et al., 2020) and consistent with current autism diagnostic criteria (WHO, 2022), autism traits appeared to be weakly related to language and NVIQ. Thus, use of autism trait measures to make inferences about broader abilities, such as structural language, may not be motivated. Descriptively, participants showed different patterns

from Manenti et al. (2024), where no autistic adults showed high NVIQ and LI. Here, two participants with high NVIQ met LI criteria at < -1.25 *SD* and < -1.5 *SD* on 2 or more measures: sentence production, sentence repetition, Semantic Relationships, and Understanding Spoken Paragraphs. While the first three measures are sensitive to LI in nonautistic youth (Calder et al., 2023), their clinical utility could differ for older autistic individuals, as tasks like sentence production require pragmatic skills (Wiig et al., 2013b). Also, some participants scored in the typical range on four or more measures and had low NVIQ (n = 2) or borderline NVIQ (n = 12), meaning that LI classification depends on the exact measures.

Overall, findings support heterogeneity in clinical classification outcomes for LI in autistic adolescents and young adults varying in levels of autism traits and NVIQ. Descriptive evaluation of assessment performance and clusters suggest that measures were sensitive to variability, but the ability to test hypotheses is limited. Variability in assessment profiles indicate dissociation of language and NVIQ, which could indicate modularity of language. Yet given discrepancy from prior work showing variable – and sometimes – higher LI classification rates, study findings could be an artifact of measurement limitations.

Implications for Evidence-Based Practice

Study findings also have implications for practice (Sackett et al., 1996), including practitioners and researchers who interpret and use assessments to build the evidence base (Durkin et al., 2015). Unlike prior work that has focused on sentence repetition and nonword repetition to screen LI in autistic youth and adults, this study included a wider range of measures and found significant heterogeneity. Clinical classification varied by definition, severity of LI cutoff, and by measure. Contrary to expectations, single-domain receptive vocabulary and nonword repetition each had two discrepant outcomes from other measures at -1 *SD* and -1.25

SD. Further, sentence repetition had more discrepant outcomes at -1 *SD* (*n* = 3)—which was the same as single-word expressive vocabulary—than at -1.25 *SD* or -1.5 *SD* (*n* = 1). These findings caution against relying on a single measure to make inferences about language skills, particularly in autistic individuals who are the most likely to lack appropriate care (Weitlauf et al., 2024).

Findings also offer information for autistic individuals and their support networks navigating services. Assessment is one step in demonstrating a need for services and identifying areas where autistic people—especially adults—may want support, such as communication (Cummins et al., 2020). Amid a lack of evidence on language in autism beyond childhood (Howlin & Taylor, 2015), findings provide descriptive evidence to facilitate understanding assessment performance and making decisions about services. For example, if eligibility criteria for speech-language services include a cutoff of -1.5 *SD* on a given number of language measures, a client could advocate for measures that are less likely to lead to discrepant outcomes (e.g., recalling sentences versus expressive vocabulary). Given dissociation of NVIQ and language and of autism traits from NVIQ and language, which are consistent with published patterns (Manenti et al., 2024; Silleresi et al., 2020), a client could advocate for evaluation based on precise assessment versus assumptions (Kover & Abbeduto, 2023).

Limitations

Although this study was exploratory, it had limitations. This study was cross-sectional and had a small sample. When coupled with data-driven analyses like clustering, findings are highly sample dependent and may not replicate (Cox & Sosine, 2023). Lack of normed overall expressive-receptive language measures for participants over age 21 also limited interpretation of findings. While use of age 21 norms in this study aligned with previous work (Botting, 2020; Clegg et al., 2021; Fidler et al., 2011; Poll et al., 2010), measures normed for adults are needed

to address questions, such as how hallmarks of LI (e.g., morphosyntax) manifest beyond childhood (Eigsti et al., 2011; Tomblin, 2008). This study included participants who use primarily spoken language to communicate, which hinders generalizability. Autistic adolescents and young adults who use minimal spoken language have clinically relevant variability in their language skills (Bal et al., 2016; Butler et al., 2023), but language assessments may not be designed for all communication modalities. Relatedly, this study used direct clinical language assessment, which limits ecological validity, especially as language assessment outcomes in autism can vary by context and setting (Barokova et al., 2020; Butler et al., 2022).

Future Directions

Limitations also lay groundwork for future directions. To better understand language across the autism spectrum, research is needed to understand how to assess and characterize language skills beyond a label (Schaeffer et al., 2023). To support holistic approaches to assessment, future work with transition-aged autistic individuals could compare clinical assessment outcomes with clinical judgment fay(Johnson et al., 1999a). Similarly, given the role of communicative feedback in language learning (Nikolaus & Fourtassi, 2023), comparison of clinical assessment with real-world language use is needed. Such investigations could consider the contributions of both autistic individuals and communication partners in successful language use in naturalistic settings (Crompton et al., 2020). To test hypotheses about language in autism during the transition to adulthood, such as to what extent language and cognitive profiles vary (WHO, 2022), robust sampling and big data analysis to support gold-standard practices (e.g., a hold-out sample) are needed (Lombardo et al., 2019). All of these directions align with precision-driven approaches to developing supports that support the goals and meet the needs of autistic individuals throughout the lifespan.

Conclusion

In evaluating language assessment across linguistic domains in autistic adolescents and young adults, this study showed that clinical classification of LI differed by operational definition, measure, and cutoff severity. Levels of language and NVIQ dissociated in approximately one-third of the sample. Exploratory analysis indicated heterogeneity in sample performance, even within participants grouped on the basis of language and NVIQ levels. Overall, findings support use of multi-domain, multi-measure language assessment and a need for adult clinical language measures.

References

Abdi, H., & Williams, L. J. (2010). Principal component analysis. *WIREs Computational Statistics*, *2*(4), 433-459. https://doi.org/10.1002/wics.101

American Association on Intellectual and Developmental Disabilities. (2024). *Defining criteria for intellectual disability*. Author. Retrieved June 2 from https://www.aaidd.org/intellectual-disability/definition

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders: DSM-5* (5th ed.). Author.

Anderson, C., & Butt, C. (2018). Young adults on the autism spectrum: The struggle for appropriate services. *Journal of Autism and Developmental Disorders*, *48*(11), 3912-3925. https://doi.org/10.1007/s10803-018-3673-z

Arunachalam, S., Avtushka, V., Luyster, R. J., & Guthrie, W. (2022). Consistency and inconsistency in caregiver reporting of vocabulary. *Language Learning and Development*, *18*(1), 81-96. https://doi.org/10.1080/15475441.2021.1931233

Arunachalam, S., & Luyster, R. J. (2016). The integrity of lexical acquisition mechanisms in autism spectrum disorders: A research review. *Autism Research*, *9*(8), 810-828. https://doi.org/https://doi.org/10.1002/aur.1590

Baird, G., Simonoff, E., Pickles, A., Chandler, S., Loucas, T., Meldrum, D., & Charman, T. (2006). Prevalence of disorders of the autism spectrum in a population cohort of children in South Thames: the Special Needs and Autism Project (SNAP). *The Lancet*, *368*(9531), 210-215. https://doi.org/10.1097/DBP.0b013e3181d5d03b

Baixauli, I., Colomer, C., Roselló, B., & Miranda, A. (2016). Narratives of children with high-functioning autism spectrum disorder: A meta-analysis. *Research in Developmental Disabilities*, *59*, 234-254. https://doi.org/10.1016/j.ridd.2016.09.007

Bal, V. H., Katz, T., Bishop, S. L., & Krasileva, K. (2016). Understanding definitions of minimally verbal across instruments: Evidence for subgroups within minimally verbal children and adolescents with autism spectrum disorder. *Journal of Child Psychology and Psychiatry*, *57*(12), 1424-1433. https://doi.org/10.1111/jcpp.12609

Barokova, M. D., Hassan, S., Lee, C., Xu, M., & Tager-Flusberg, H. (2020). A comparison of natural language samples collected from minimally and low-verbal children and adolescents with autism by parents and examiners. *Journal of Speech, Language, and Hearing Research*, *63*(12), 4018-4028. https://doi.org/10.1044/2020_jslhr-20-00343

Bartlett, M. S. (1951). The effect of standardization on a Chi-square approximation in factor analysis. *Biometrika*, *38*, 337-344.

Benton, A., & Hamsher, K. (1978). *Multilingual Aphasia Examination*. AJA Associates.

Bishop, D. V., Snowling, M. J., Thompson, P. A., Greenhalgh, T., Consortium, C., Adams, C., Archibald, L., Baird, G., Bauer, A., & Bellair, J. (2017). Phase 2 of CATALISE: A multinational and multidisciplinary Delphi consensus study of problems with language development: Terminology. *Journal of Child Psychology and Psychiatry*, *58*(10), 1068-1080. https://doi.org/10.1111/jcpp.12721

Bishop, S. L., & Lord, C. (2023). Commentary: Best practices and processes for assessment of autism spectrum disorder–the intended role of standardized diagnostic instruments. *Journal of Child Psychology and Psychiatry*, *64*(5), 834-838. https://doi.org/10.1111/jcpp.13802

Bonferroni, C. (1936). Teoria statistica delle classi e calcolo delle probabilita. *Pubblicazioni del R Istituto Superiore di Scienze Economiche e Commericiali di Firenze*, *8*, 3-62.

Botting, N. (2020). Language, literacy and cognitive skills of young adults with developmental language disorder (DLD). *International Journal of Language & Communication Disorders*, *55*(2), 255-265. https://doi.org/10.1111/1460-6984.12518

Boucher, J. (2012). Research review: structural language in autistic spectrum disorder–characteristics and causes. *Journal of Child Psychology and Psychiatry*, *53*(3), 219-233. https://doi.org/10.1111/j.1469-7610.2011.02508.x

Bouguettaya, A., Yu, Q., Liu, X., Zhou, X., & Song, A. (2015). Efficient agglomerative hierarchical clustering. *Expert Systems with Applications*, *42*(5), 2785-2797. https://doi.org/10.1016/j.eswa.2014.09.054

Brignell, A., Morgan, A. T., Woolfenden, S., Klopper, F., May, T., Sarkozy, V., & Williams, K. (2018). A systematic review and meta-analysis of the prognosis of language outcomes for individuals with autism spectrum disorder. *Autism & Developmental Language Impairments*, *3*, 2396941518767610. https://doi.org/10.1177/2396941518767610

Bryan, K. L. (1989). *The Right Hemisphere Language Battery*. Far Communications.

Burke, M. M., Cheung, W. C., Best, M., DaWalt, L. S., & Taylor, J. L. (2024). Measuring what matters: Considerations for the measurement of services for individuals with autism. *Journal of Developmental and Physical Disabilities*, *36*, 423-439. https://doi.org/10.1007/s10882-023-09916-6

Butler, L. K., LaValle, C., Shen, L. S., Palana, J., Schwartz, S., Liu, C., Peterman, N., & Tager-Flusberg, H. (2022). Remote natural language sampling of parents and children with autism spectrum disorder: Role of activity and language level. *Frontiers in Communication,*, *7*, 820564. https://doi.org/10.3389/fcomm.2022.820564

Butler, L. K., Shen, L., Chenausky, K. V., La Valle, C., Schwartz, S., & Tager-Flusberg, H. (2023). Lexical and morphosyntactic profiles of autistic youth with minimal or low spoken language skills. *American Journal of Speech-Language Pathology*, *32*(2), 733-747. https://doi.org/10.1044/2022_AJSLP-22-00098

Calder, S. D., Brennan-Jones, C. G., Robinson, M., Whitehouse, A., & Hill, E. (2023). How we measure language skills of children at scale: A call to move beyond domain-specific tests

as a proxy for language. *International Journal of Speech-Language Pathology*, *32*(2), 440-448. https://doi.org/10.1080/17549507.2023.2171488

Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate behavioral research*, *1*(2), 245-276. https://doi.org/10.1207/s15327906mbr0102_10

Chatham, C., Taylor, K., Charman, T., Liogier D'Ardhuy, X., Eule, E., Fedele, A., Hardan, A., Loth, E., Murtagh, L., & del Valle Rubido, M. (2018). Adaptive behavior in autism: Minimal clinically important differences on the Vineland□II. *Autism Research*, *11*(2), 270-283. https://doi.org/10.1002/aur.1874

Clegg, J., Crawford, E., Spencer, S., & Matthews, D. (2021). Developmental Language Disorder (DLD) in young people leaving care in England: A study profiling the language, literacy and communication abilities of young people transitioning from care to independence. *International Journal of Environmental Research and Public Health*, *18*(8), 4107. https://doi.org/10.3390/ijerph18084107

Cochran, W. G. (1950). The comparison of percentages in matched samples. *Biometrika*, *37*(3/4), 256-266.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, *20*(1), 37-46. https://doi.org/10.1177/001316446002000104

Constantino, J. N. (2012). *Social Responsiveness Scale, Second Edition*. W. P. Services.

Constantino, J. N., & Grubler, C. P. (2012). *Social Responsiveness Scale-Second Ed.* Western Psychological Services.

Cox, D., & Sosine, J. (2023). Influence of sample size, feature set, and algorithm on cluster analyses for patients with autism spectrum disorders. *PsyArXiv*. https://doi.org/10.21203/rs.3.rs-3351792/v1

Creemers, A., Schaeffer, Jeannette. (2015). Specific language impairment and high functioning autism: Evidence for distinct etiologies and for modularity of grammar and pragmatics. 6th Conference on Generative Approaches to Language Acquisition North America (GALANA), Somerville, MA.

Crompton, C. J., Ropar, D., Evans-Williams, C. V., Flynn, E. G., & Fletcher-Watson, S. (2020). Autistic peer-to-peer information transfer is highly effective. *Autism*, *24*(7), 1704-1712. https://doi.org/10.1177/1362361320919286

Cummins, C., Pellicano, E., & Crane, L. (2020). Autistic adults' views of their communication skills and needs. *International Journal of Language & Communication Disorders*, *55*(5), 678-689. https://doi.org/10.1111/1460-6984.12552

Dalmaijer, E. S., Nord C. L., & Astle D. E. (2022). Statistical power for cluster analysis. *BMC Bioinformatics*, *23*(1), 205. https://doi.org/10.1186/s12859-022-04675-1

Day, W. H., & Edelsbrunner, H. (1984). Efficient algorithms for agglomerative hierarchical clustering methods. *Journal of Classification*, *1*(1), 7-24. https://doi.org/10.1007/BF01890115

Dudoit, S., & Fridlyand, J. (2002). A prediction-based resampling method for estimating the number of clusters in a dataset. *Genome Biology*, *3*(7), research0036.0031. https://doi.org/10.1186/gb-2002-3-7-research0036

Dunn, D. M. (2019). *Peabody Picture Vocabulary Test, Fifth Edition: Manual*. Pearson.

Dunn, L. M., & Dunn, L. M. (1981). *Peabody Picture Vocabulary Test-Revised*. American Guidance Service, Inc.

Dunn, L. M., Dunn, L. M., & Whetton, C. (1982). *The British Picture Vocabulary Scale: Manual for the short and long forms*. NFER-Nelson.

Dunn, O. J. (1964). Multiple comparisons using rank sums. *Technometrics*, *6*(3), 241-252. https://doi.org/10.1080/00401706.1964.10490181

Durkin, M. S., Elsabbagh, M., Barbaro, J., Gladstone, M., Happe, F., Hoekstra, R. A., Lee, L. C., Rattazzi, A., Stapel Wax, J., & Stone, W. L. (2015). Autism screening and diagnosis in low resource settings: challenges and opportunities to enhance research and services worldwide. *Autism Research*, *8*(5), 473-476. https://doi.org/10.1002/aur.1575

Eigsti, I.-M., de Marchena, A. B., Schuh, J. M., & Kelley, E. (2011). Language acquisition in autism spectrum disorders: A developmental review. *Research in Autism Spectrum Disorders*, *5*(2), 681-691.

Eilenberg, J. S., Paff, M., Harrison, A. J., & Long, K. A. (2019). Disparities based on race, ethnicity, and socioeconomic status over the transition to adulthood among adolescents and young adults on the autism spectrum: A systematic review. *Current Psychiatry Reports*, *21*(5), 32. https://doi.org/10.1007/s11920-019-1016-1

Every Student Succeeds Act, Pub. L. No. 114-95, § 1001-1605 (2015).

Ferré, S., Mahé, O., Manenti, M., Pimenta, H., & Prévost, P. (2022). Detecting a phonological deficit in adults-an adult module for LITMUS-QU-NWR-FR. XXXIVe Journées d'Études sur la Parole--JEP 2022,

Fidler, L. J., Plante, E., & Vance, R. (2011). Identification of adults with developmental language impairments. *American Journal of Speech-Language Pathology*, *20*(1), 2-13. https://doi.org/10.1044/1058-0360(2010/09-0096)

Fidler, L. J., Plante, E., & Vance, R. (2013). Erratum. *American Journal of Speech-Language Pathology*, *22*(3), 577-577. https://doi.org/10.1044/1058-0360(2013/13-0018)

Gaeta, L., & Brydges, C. R. (2020). An examination of effect sizes and statistical power in speech, language, and hearing research. *Journal of Speech, Language, and Hearing Research*, *63*(5), 1572-1580. https://doi.org/10.1044/2020_JSLHR-19-00299

Gardner, M. F. (1982). *Expressive One Word Picture Vocabulary Test (Upper Extensions)*. Western Psychological Services.

Girolamo, T., Ghali, S., Campos, I., & Ford, A. (2022). Interpretation and use of language assessments for diverse school-age individuals. *Perspectives of the ASHA special interest groups*. https://doi.org/10.1044/2022_PERSP-21-00322

Girolamo, T., & Rice, M. L. (2022). Language impairment in autistic young adults. *Journal of Speech, Language, and Hearing Research*. https://doi.org/10.1044/2022_JSLHR-21-00517

Girolamo, T., Shen, L., Monroe Gulick, A., Rice, M. L., & Eigsti, I.-M. (2023). Studies assessing domains pertaining to structural language in autism vary in reporting practices and approaches to assessment: A systematic review. *Autism*, 13623613231216155. https://doi.org/10.1177/13623613231216155

Grondhuis, S. N., Lecavalier, L., Arnold, L. E., Handen, B. L., Scahill, L., McDougle, C. J., & Aman, M. G. (2018). Differences in verbal and nonverbal IQ test scores in children with autism spectrum disorder. *Research in Autism Spectrum Disorders*, *49*, 47-55. https://doi.org/10.1016/j.rasd.2018.02.001

Hammill, D. D., Brown, V. L., Larsen, S. C., & Wiederholt, J. L. (1994). *Test of Adolescent/Adult Language-3*. Pro-Ed.

Hammill, D. D., Brown, V. L., Larsen, S. C., & Wiederholt, J. L. (2007). *Test of Adolescent and Adult Language-Fourth Edition (TOAL-4)*. Pro-Ed.

Hartigan, J. A., & Wong, M. A. (1979). Algorithm AS 136: A k-means clustering algorithm. *Journal of the royal statistical society. series c (applied statistics)*, *28*(1), 100-108. https://doi.org/10.2307.2346830

Howlin, P., Goode, S., Hutton, J., & Rutter, M. (2004). Adult outcome for children with autism. *Journal of Child Psychology and Psychiatry*, *45*(2), 212-229. https://doi.org/10.1111/j.1469-7610.2004.00215.x

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*, *55*(1), 49-58. https://doi.org/10.1111/jcpp.12115

Howlin, P., & Taylor, J. L. (2015). Addressing the need for high quality research on autism in adulthood. *Autism*, *19*(7), 771-773. https://doi.org/10.1177/1362361315595582

Husson, F., Josse, J., & Pages, J. (2010). *Principal component methods-hierarchical clustering-partitional clustering: Why would we need to choose for visualizing data* (Technical Report of the Applied Mathematics Department, Issue. Agrocampus Ouest.

IBM Corp. (2025). *IBM SPSS Statistics for Macintosh*. In (Version 31.0.0.0) IBM Corp.

Individuals with Disabilities Education Act, U.S.C. § 1400 (2004).

Individuals with Disabilities Education Improvement Act of 2004, 20 U.S.C. § 1400 et seq. (2018).

Johnson, C. J., Beitchman, J. H., Young, A., Escobar, M., Atkinson, L., Wilson, B., Brownlie, E. B., Douglas, L., Taback, N., & Lam, I. (1999a). Fourteen-year follow-up of children with and without speech/language impairments: Speech/language stability and outcomes. *Journal of Speech, Language, and Hearing Research*, *42*(3), 744-760. https://doi.org/10.1044/jslhr.4203.744

Johnson, C. J., Taback, N., Escobar, M., Wilson, B., & Beitchman, J. (1999b). Local norming of the Test of Adolescent/Adult Language-3 in the Ottawa Speech and Language Study. *Journal of Speech, Language, and Hearing Research*, *42*(3), 761-766. https://doi.org/10.1044/jslhr.4203.761

Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika*, *39*(1), 31-36. https://doi.org/10.1007/BF02291575

Kertesz, A. (1982). *Western Aphasia Battery*. Grune and Stratton.

Kover, S. T., & Abbeduto, L. (2023). Toward equity in research on intellectual and developmental disabilities. *American Journal on Intellectual and Developmental Disabilities*, *128*(5), 350-370. https://doi.org/10.1352/1944-7558-128.5.350

Levy, S. E., Giarelli, E., Lee, L.-C., Schieve, L. A., Kirby, R. S., Cunniff, C., Nicholas, J., Reaven, J., & Rice, C. E. (2010). Autism spectrum disorder and co-occurring developmental, psychiatric, and medical conditions among children in multiple populations of the United States. *Journal of Developmental & Behavioral Pediatrics*, *31*(4), 267-275. https://doi.org/10.1097/DBP.0b013e3181d5d03b

Lewis, F. M., Woodyatt, G. C., & Murdoch, B. E. (2008). Linguistic and pragmatic language skills in adults with autism spectrum disorder: A pilot study. *Research in Autism Spectrum Disorders*, *2*(1), 176-187. https://doi.org/10.1016/j.rasd.2007.05.002

Little, R. J., & Rubin, D. B. (2019). *Statistical analysis with missing data* (Vol. 793). John Wiley & Sons.

Lombardo, M. V., Lai, M.-C., & Baron-Cohen, S. (2019). Big data approaches to decomposing heterogeneity across the autism spectrum. *Molecular Psychiatry*, *24*(10), 1435-1450. https://doi.org/10.1038/s41380-018-0321-0

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

Loucas, T., Charman, T., Pickles, A., Simonoff, E., Chandler, S., Meldrum, D., & Baird, G. (2008). Autistic symptomatology and language ability in autism spectrum disorder and specific language impairment. *Journal of Child Psychology and Psychiatry*, *49*(11), 1184-1192. https://doi.org/10.1111/j.1469-7610.2008.01951.x

Lu, J.-F., Tang, J., Tang, Z.-M., & Yang, J.-Y. (2008). Hierarchical initialization approach for K-Means clustering. *Pattern Recognition Letters*, *29*(6), 787-795. https://doi.org/10.1016/j.patrec.2007.12.009

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. https://doi.org/10.1016/j.cpr.2013.11.002

Manenti, M., Ferré, S., Tuller, L., Houy-Durand, E., Bonnet-Brilhault, F., & Prévost, P. (2024). Profiles of structural language and nonverbal intellectual abilities in verbal autistic adults. *Research in Autism Spectrum Disorders*, *114*, 102361. https://doi.org/10.1016/j.lingua.2023.103598

Manenti, M., Tuller, L., Houy-Durand, E., Bonnet-Brilhault, F., & Prévost, P. (2023). Assessing structural language skills of autistic adults: Focus on sentence repetition. *Lingua*, *294*, 103598. https://doi.org/10.1016/j.rasd.2024.102361

Mawhood, L., Howlin, P., & Rutter, M. (2000). Autism and developmental receptive language disorder—A comparative follow-up in early adult life. I: Cognitive and language outcomes. *The Journal of Child Psychology and Psychiatry and Allied Disciplines*, *41*(5), 547-559. https://doi.org/10.1111/1469-7610.00642

Morice, R., & McNicol, D. (1985). The comprehension and production of complex syntax in schizophrenia. *Cortex*, *21*(4), 567-580. https://doi.org/10.1016/S0010-9452(58)80005-2

Nikolaus, M., & Fourtassi, A. (2023). Communicative feedback in language acquisition. *New Ideas in Psychology*, *68*, 100985. https://doi.org/10.1016/j.newideapsych.2022.100985

Nitido, H., & Plante, E. (2020). Diagnosis of developmental language disorder in research studies. *Journal of Speech, Language, and Hearing Research*, *63*(8), 2777-2778. https://doi.org/10.1044/2020_JSLHR-20-00091

Norbury, C. F., Gooch, D., Wray, C., Baird, G., Charman, T., Simonoff, E., Vamvakas, G., & Pickles, A. (2016). The impact of nonverbal ability on prevalence and clinical presentation of language disorder: Evidence from a population study. *Journal of Child Psychology and Psychiatry*, *57*(11), 1247-1257. https://doi.org/10.1111/jcpp.12573

Oetting, J. B., & McDonald, J. L. (2002). Methods for characterizing participants' nonmainstream dialect use in child language research. *Journal of Speech, Language & Hearing Research*, *45*(3), 505-518. https://doi.org/10.1044/1092-4388(2002/040)

Office of Management and Budget. (1997). *Revisions to the standards for the classification of federal data on race and ethnicity* (Federal Register Notice, Issue. https://obamawhitehouse.archives.gov/omb/fedreg_1997standards

Orlov, K. (2024). *KO Macros*. In *Kirill's SPSS Macros*, Raynald's SPSS Tools,. https://www.spsstools.net/en/macros/KO-spssmacros/

Pearson. (2025). *Staying connected through telepractice*. Retrieved January 16 from https://www.pearsonassessments.com/professional-assessments/digital-solutions/telepractice/about.html

Pizzano, M., Shire, S., Shih, W., Levato, L., Landa, R., Lord, C., Smith, T., & Kasari, C. (2024). Profiles of minimally verbal autistic children: Illuminating the neglected end of the spectrum. *Autism Research*, *n/a*(n/a). https://doi.org/https://doi.org/10.1002/aur.3151

Poll, G. H., Betz, S. K., & Miller, C. A. (2010). Identification of clinical markers of specific language impairment in adults. *Journal of Speech, Language & Hearing Research*, *53*, 414-429. https://doi.org/10.1044/1092-4388(2009/08-0016)

Raven, J., Rust, J., Chan, F., & Zhou, X. (2018). *Raven's 2 Progressive Matrices, Clinical Edition*. Pearson.

Rosen, N. E., Lord, C., & Volkmar, F. R. (2021). The diagnosis of autism: From Kanner to DSM-III to DSM-5 and beyond. *Journal of Autism and Developmental Disorders*, *51*(12), 4253-4270. https://doi.org/10.1007/s10803-021-04904-1

Roux, A. M., Chvasta, K., McLean, K. J., Carey, M., Perez Liz, G., Tomczuk, L., Lopez, K., Assing-Murray, E., Shattuck, P. T., & Shea, L. L. (2024). Challenges and opportunities in transitioning autistic individuals Into adulthood. *Pediatrics*, *154*(5), e2024067195. https://doi.org/10.1542/peds.2024-067195

Russell, G., Mandy, W., Elliott, D., White, R., Pittwood, T., & Ford, T. (2019). Selection bias on intellectual ability in autism research: A cross-sectional review and meta-analysis. *Molecular Autism*, *10*(1), 1-10. https://doi.org/10.1186/s13229-019-0260-x

Sackett, D. L., Rosenberg, W. M., Gray, J. M., Haynes, R. B., & Richardson, W. S. (1996). Evidence based medicine: What it is and what it isn't. In (Vol. 312, pp. 71-72): BMJ.

Schaeffer, J., Abd El-Raziq, M., Castroviejo, E., Durrleman, S., Ferré, S., Grama, I., Hendriks, P., Kissine, M., Manenti, M., Marinis, T., Meir, N., Novogrodsky, R., Perovic, A., Panzeri, F., Silleresi, S., Sukenik, N., Vicente, A., Zebib, R., Prévost, P., & Tuller, L. (2023). Language in autism: Domains, profiles and co-occurring conditions. *Journal of Neural Transmission*, *130*(3), 433-457. https://doi.org/10.1007/s00702-023-02592-y

Selin, C. M., Rice, M. L., Girolamo, T. M., & Wang, C. J. (2022). Work setting effects on speech-language pathology practice: Implications for Identification of children with specific language impairment. *American Journal of Speech-Language Pathology*, *31*(2), 854-880. https://doi.org/10.1044/2021_AJSLP-21-00024

Semel E, Wiig E.H., & Secord, W. A. (2003). *Clinical Evaluation of Language Fundamentals-4th ed.* Harcourt.

Semel, E., Wiig, E. H., & Secord, W. (2017). *Clinical Evaluation of Language Fundamentals—Fifth Edition (CELF-5 UK)*. Pearson.

Semel, E. M., Wiig E.H., & Secord, W. (2006). *Clinical Evaluation of Language Fundamentals-4th ed.* Psych-Corp.

Shaw, K. A., Williams, S., Patrick, M. E., Valencia-Prado, M., Durkin, M. S., Howerton, E. M., Ladd-Acosta, C. M., Pas, E. T., Bakian, A. V., & Bartholomew, P. (2025). Prevalence and early identification of autism spectrum disorder among children aged 4 and 8 years-Autism and Developmental Disabilities Monitoring Network, 16 Sites, United States, 2022. *MMWR Surveillance Summaries*, *74*(2), 1-22. https://doi.org/10.15585/mmwr.ss7402a1

Shriberg, L. D., Lohmeier, H. L., Campbell, T. F., Dollaghan, C. A., Green, J. R., & Moore, C. A. (2009a). A nonword repetition task for speakers with misarticulations: The Syllable Repetition Task (SRT). https://doi.org/10.1044/1092-4388(2009/08-0047)

Shriberg, L. D., Lohmeier, H. L., Campbell, T. F., Dollaghan, C. A., Green, J. R., & Moore, C. A. (2009b). A nonword repetition task for speakers with misarticulations: The Syllable

Repetition Task (SRT). *Journal of Speech, Language & Hearing Research,, 52*(5), 1189-1212. https://doi.org/10.1044/1092-4388(2009/08-0047)

Shriberg, L. D., & Mabie, H. L. (2017). *Speech and motor speech assessment findings in eight complex neurodevelopmental disorders* (24). www.waisman.wisc.edu/phonology/techreports/TREP24.PDF

Silleresi, S., Prévost, P., Zebib, R., Bonnet□Brilhault, F., Conte, D., & Tuller, L. (2020). Identifying language and cognitive profiles in children with ASD via a cluster analysis exploration: Implications for the new ICD□11. *Autism Research*, *13*(7), 1155-1167. https://doi.org/10.1002/aur.2268

Tomblin, J. B. (1996). The big picture of SLI: Epidemiology of SLI. Symposium on Research in Child Language Disorders, Madison, WI,

Tomblin, J. B. (2008). Validating diagnostic standards for specific language impairment using adolescent outcomes. In C. F. Norbury, J. B. Tomblin, & D. V. M. Bishop (Eds.), *Understanding developmental language disorders: From theory to practice* (pp. 93-116). Psychology Press.

Tomblin, J. B., Records, N. L., & Zhang, X. (1996). A system for the diagnosis of specific language impairment in kindergarten children. *Journal of Speech, Language, and Hearing Research*, *39*(6), 1284-1294. https://doi.org/10.1044/jshr.3906.1284

Tuller, L., Konrad, I., Ridet, J., & Prévost, P. (2021). *LITMUS-SR-FR-Adults*. University of Tours.

Ward Jr, J. H. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association*, *58*(301), 236-244. https://doi.org/10.1080/01621459.1963.10500845

Wechsler, D. (2008). *Wechsler Adult Intelligence Scales-4th Ed.* The Psychological Corporation.

Weir, E., Allison, C., & Baron-Cohen, S. (2022). Autistic adults have poorer quality healthcare and worse health based on self-report data. *Molecular Autism*, *13*(1), 23. https://doi.org/10.1186/s13229-022-00501-w

Weitlauf, A. S., Miceli, A., Vehorn, A., Dada, Y., Pinnock, T., Harris, J. W., Hine, J., & Warren, Z. (2024). Screening, diagnosis, and intervention for autism: Experiences of Black and multiracial families seeking care. *Journal of Autism and Developmental Disorders*, *54*(3), 931-942. https://doi.org/10.1007/s10803-022-05861-z

Wiig, E. H., Semel, E., & Secord, W. (2013a). *Clinical Evaluation of Language Fundamentals-5th Ed.*

Wiig, E. H., Semel, E., & Secord, W. A. (2013b). *Clinical Evaluation of Language Fundamentals-5th Ed.* Pearson.

Williams, K. T. (2019). *Expressive Vocabulary Test-Third Edition: Manual*. Pearson.

World Health Organization. (2022). *International Classification of Diseases (ICD)-11*. https://icd.who.int/en

Figure Legend

Figure 1. Percent participants meeting language impairment cutoffs across language measures. LI = language impairment.

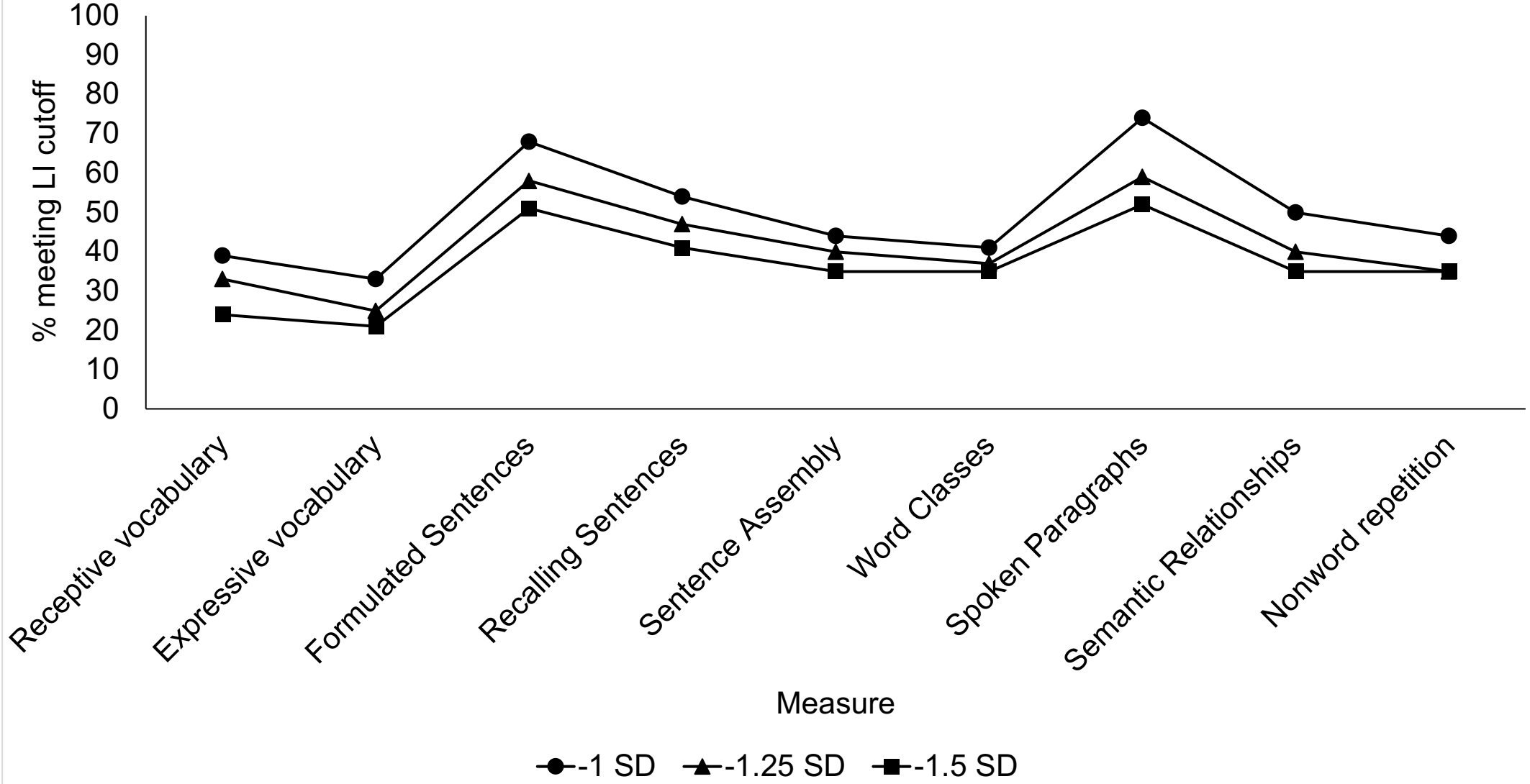
Figure 2. Frequency of significant pairwise comparisons across language measures by language impairment cutoff of -1 *SD*, -1.25 *SD*, or -1.5 *SD*

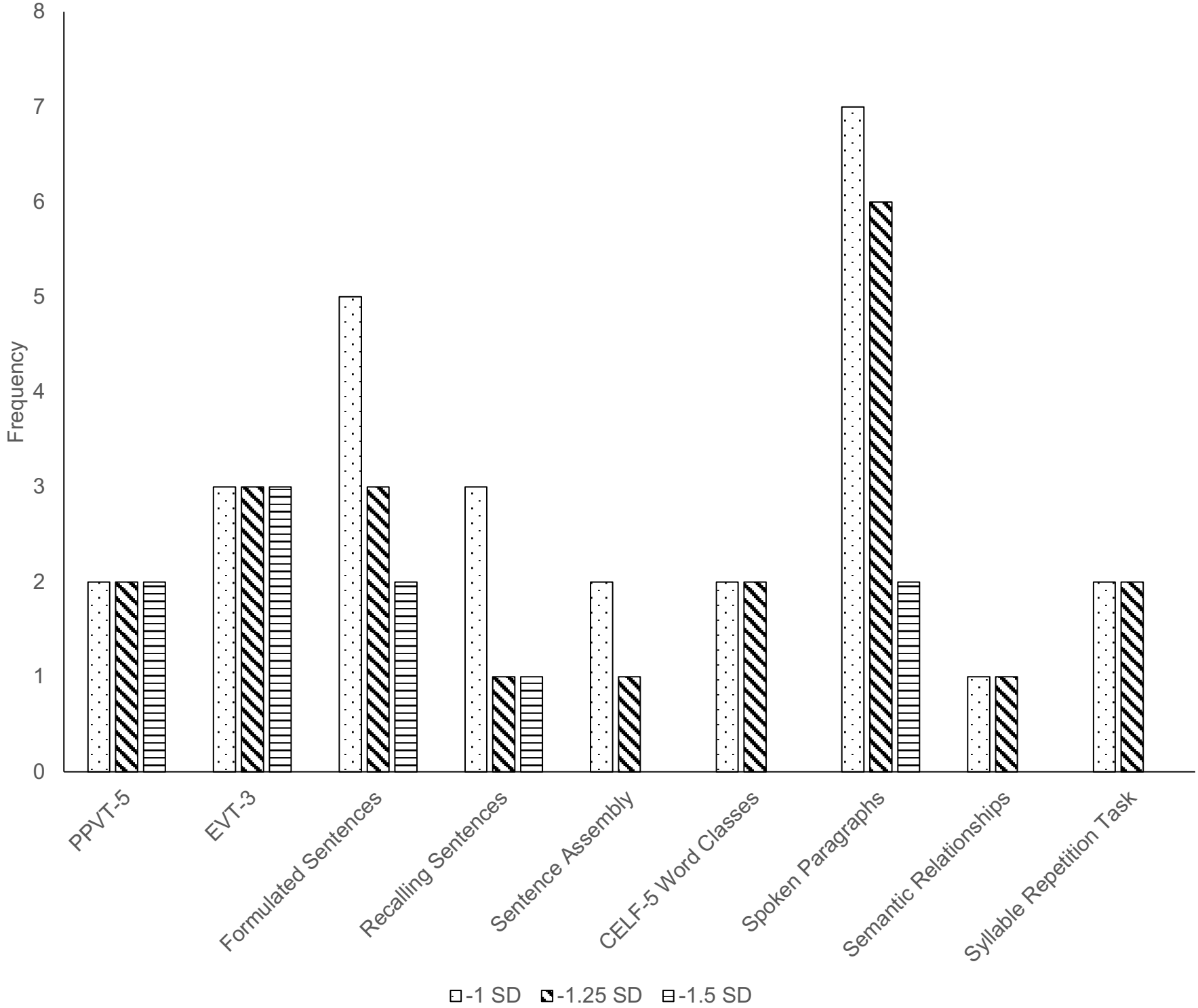
Figure 3. Scree plot and component plot in rotated space of a principal components analysis run on autism trait, language, and NVIQ *z*-scores. SCI = Social Responsiveness Scale-2 (SRS-2) social communication & social interaction, RRB = SRS-2 restricted & repetitive behaviors (Constantino, 2012). EVT3 = Expressive Vocabulary Test-3 (Williams, 2018). NVIQ = Raven's 2 nonverbal intelligence (Raven, 201 8). PPVT5 = Peabody Picture Vocabulary Test (Dunn, 2018). SRT = Syllable Repetition Task (Shriberg et al., 2009). FS = Clinical Evaluation of Language Fundamentals-5 (CELF-5) Formulated Sentences, RS = CELF-5 Recalling Sentences, CELF-5 SA = Sentence Assembly, CELF-5 SR = Semantic Relationships, CELF-5 USP = Understanding Spoken Paragraphs, WC = CELF-5 Word Classes (Wiig et al., 2013).

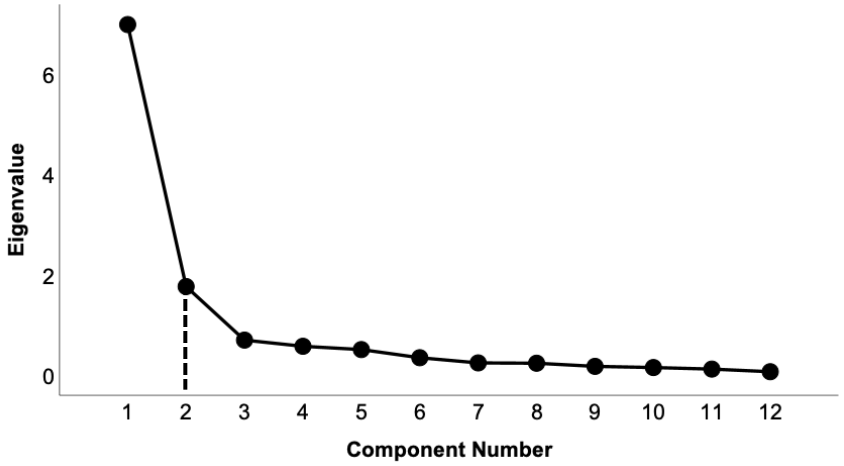
Figure 4. Dendrogram with line suggesting stopping location and scree plot

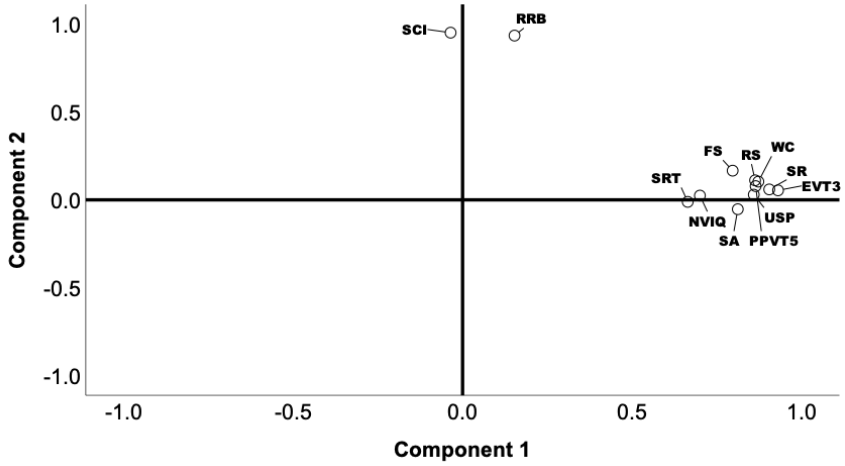
Figure 5. Clustering of participants (*N* = 74) using a *k*-means three-cluster solution

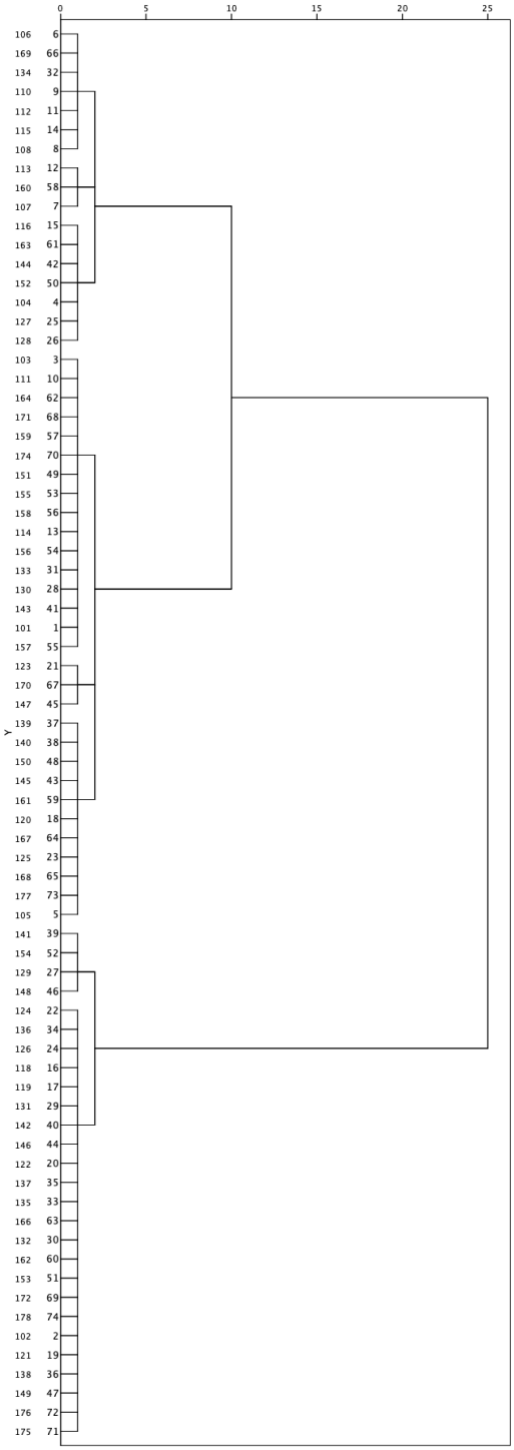
Figure 6. *Z*-scores of individual participants across language and NVIQ measures by cluster. Red = Cluster 1. Green = Cluster 2. Blue = Cluster 3. PPVT-5 = Peabody Picture Vocabulary Test-5 (Dunn, 2019). EVT-3 = Expressive Vocabulary Test-3 (Williams, 2019). FS = Formulated Sentences, RS = Recalling Sentences, SA = Sentence Assembly, WC = Word Classes, USP = Understanding Spoken Paragraphs, and SR = Semantic Relationships from the Clinical Evaluation of Language Fundamentals-5 (Wiig et al., 2013). SRT = Syllable Repetition Task (Shriberg et al., 2009). NVIQ = Raven's 2 nonverbal intelligence (Raven et al., 2018).

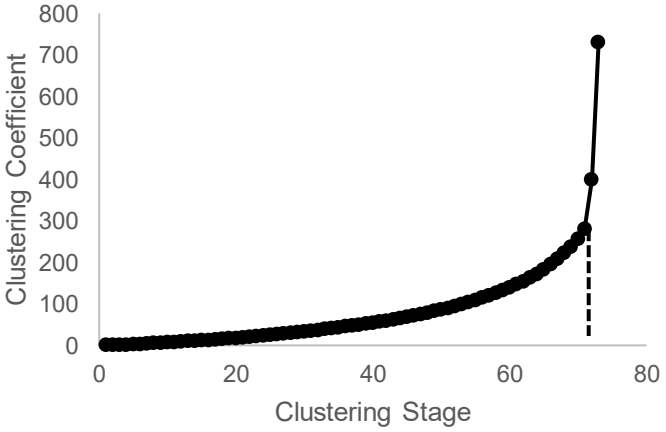

100
90
80
70
60
50
40
30
20
10
0
% meeting LI cutoff
Receptive vocabulary
Expressive vocabulary
Formulated Sentences
Recalling Sentences
Sentence Assembly
Word Classes
Spoken Paragraphs
Semantic Relationships
Nonword repetition
Measure
-1 SD
-1.25 SD
-1.5 SD

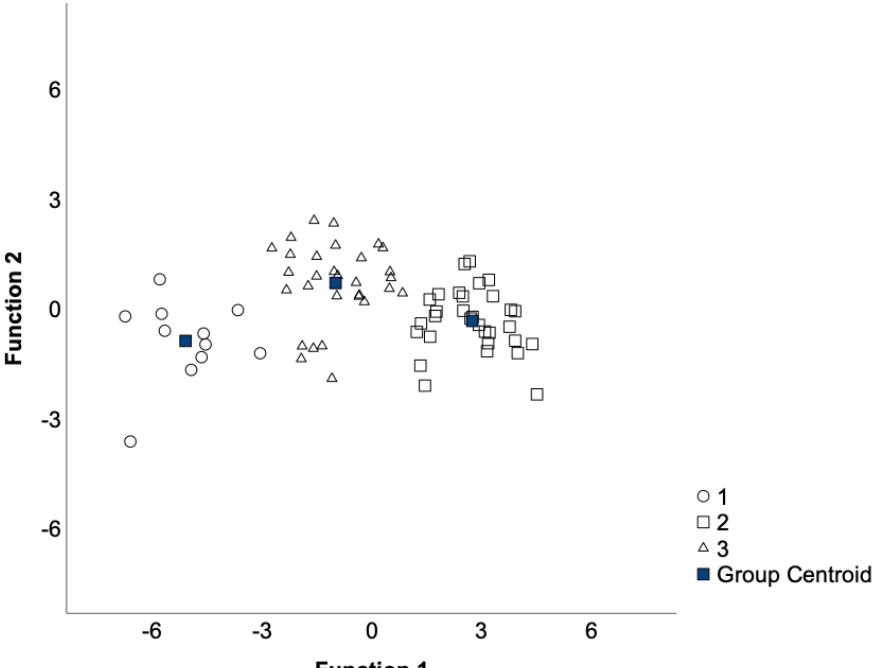

8
7
6
5
4
3
2
1
0
Frequency
PPVT-5
EVT-3
Formulated Sentences
Recalling Sentences
Sentence Assembly
CELF-5 Word Classes
Spoken Paragraphs
Semantic Relationships
Syllable Repetition Task
-1 SD
-1.25 SD
-1.5 SD

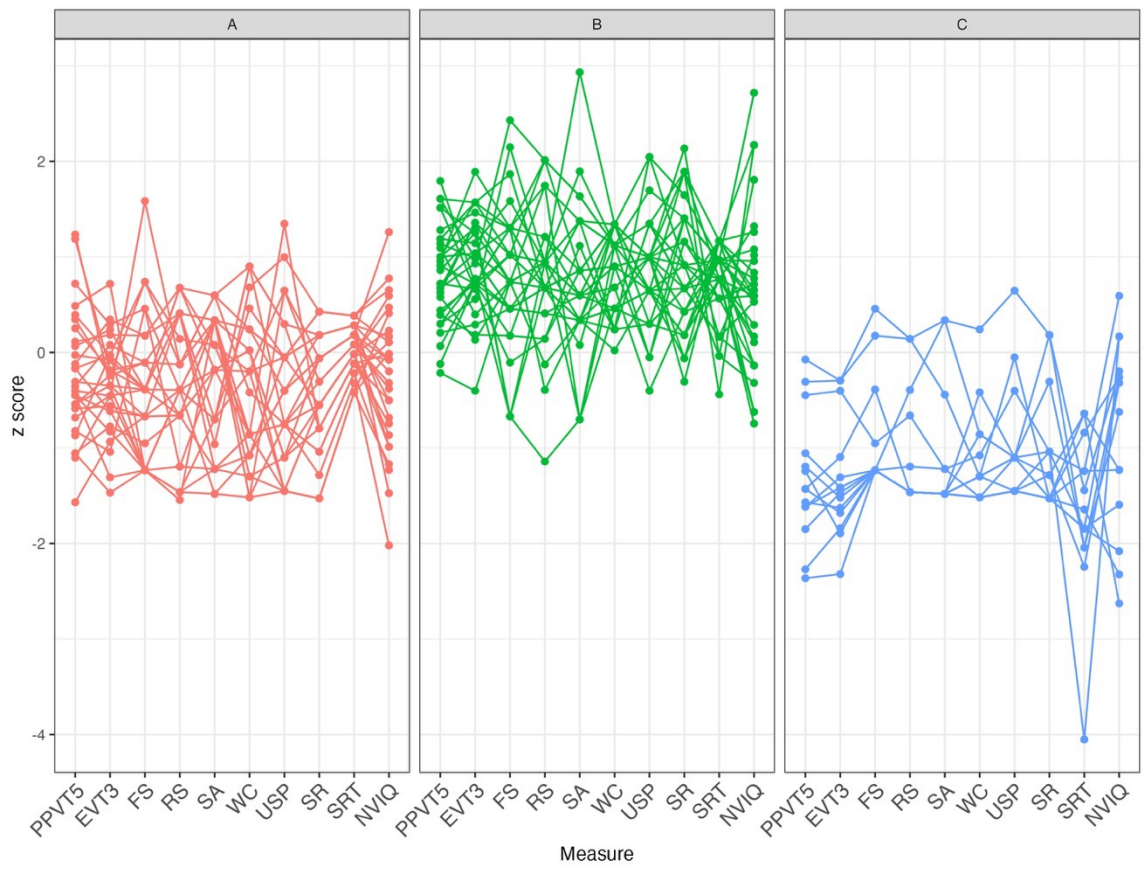

Eigenvalue
0
2
4
6
1
2
3
4
5
6
7
8
9
10
11
12
Component Number


Component 2
1.0
0.5
0.0
-0.5
-1.0
-1.0
-0.5
0.0
0.5
1.0
Component 1
SCI
RRB
FS
RS
WC
SR
EVT3
SRT
NVIQ
USP
SA
PPVT5


0
5
10
15
20
25
Y


Clustering Coefficient
800
700
600
500
400
300
200
100
0
0
20
40
60
80
Clustering Stage


6
3
0
-3
-6
Function 2
-6
-3
0
3
6
Function 1
1
2
3
Group Centroid


A
B
C
z score
2
0
-2
-4
PPVT5
EVT3
FS
RS
SA
WC
USP
SR
SRT
NVIQ
Measure